

BESZÉDKUTATÁS 2006

**TANULMÁNYOK
VÉRTES O. ANDRÁS EMLÉKÉRE**

BUDAPEST

BESZÉDKUTATÁS 2006

TANULMÁNYOK VÉRTES O. ANDRÁS EMLÉKÉRE

**Szerkesztette:
GÓSY MÁRIA**

MTA Nyelvtudományi Intézet
Kempelen Farkas Beszédkutató Laboratórium
Budapest 2006

Az egyes tanulmányokat szakterületi kompetenciával rendelkező szakemberek lektorálták.

A korpusz szerkesztésében részt vettek:

| |
|------------------|
| Gósy Mária |
| Markó Alexandra |
| Bóna Judit |
| Imre Angéla |
| Horváth Viktória |

Technikai szerkesztő: Markó Alexandra

ISSN 1218-8727

© MTA Nyelvtudományi Intézet

Felelős kiadó: Kenesei István igazgató

Készült: Regiszter Kiadó és Nyomda Kft.

TARTALOM

| | |
|---|-----|
| Gósy Mária: 95 éves lenne Vértes O. András | 5 |
| Gósy Mária: A semleges magánhangzó nyelvi funkciói | 8 |
| Nikléczy Péter: A koartikuláció a magánhangzók tiszta fázisa és az átmenetek összefüggésében | 23 |
| Olaszy Gábor: A nazálisok okozta szerkezetváltozás a zár-, rés- és zár-rés hangokban mássalhangzó-kapcsolat esetén | 32 |
| Menyhárt Krisztina: Koartikulációs folyamatok két magánhangzó kapcsolatában | 44 |
| Markó Alexandra: Nonverbális vokális jelek a társalgásban | 57 |
| Péteri Attila: Prozódia a grammatika és a pragmatika határán | 69 |
| Gráczki Tekla Etelka: Suttogott ejtésű magánhangzók észlelése | 86 |
| Bóna Judit: A megakadásjelenségek akusztikai és percepcióssajátosságai .. | 101 |
| Bata Sarolta: A meggyőzés a beszédprodukcióban és a beszédpercepcióban | 114 |
| Markó Alexandra – Bóna Judit: A spontán beszéd lejegyzésének néhány módszertani kérdése | 124 |
| Horváth Viktória: A spontán beszéd és a beszédfeldolgozás összefüggései gyerekeknél | 134 |
| Gósy Mária: A beszédhangok megkülönböztetésének fejlődése | 147 |
| Imre Angéla: Az olvasászavar és a beszédfeldolgozási folyamatok összefüggései | 160 |
| Horváth M. Szabolcs – Magony Edit – Gósy Mária: A fáradó beszéd mint betegségtünet | 172 |
| Németh Géza – Olaszy Gábor – Fék Márk: Új rendszerű, korpuszalapú gépi szövegfelolvasó fejlesztése | 183 |
| Böhm Tamás: A glottalizáció szerepe a beszélő személy felismerésében ... | 197 |
| Zainkó Csaba – Fék Márk: Beszédadatbázis prozódiajának szerepe a gépi beszéd hangzásában | 208 |
| Mihajlik Péter – Fegyő Tibor – Tatai Péter: Új eljárás a gépi beszéd-felismerés környezetfüggő beszédhangmodelljeinek kialakítására | 218 |

| | |
|--|-----|
| „Nyelvbotlás”-korpusz – 3. rész | 231 |
| 2. Morfológiai, szintaktikai hiba..... | 232 |
| 2. a) Újraindítás nélküli morfológiai változtatás..... | 232 |
| 2. b) Ismétlés toldalékjavítással..... | 232 |
| 3. Kontamináció..... | 233 |
| 4. Téves szótalálás..... | 234 |
| 4. a) Téves szókezdés..... | 236 |
| 4. b) Malapropizmus..... | 236 |
| 5. „Nyelvem hegyén van”..... | 237 |
| 6. Változtatás..... | 237 |
| 7. Újraindítás..... | 237 |
| 7. a) Újraindítás változtatással..... | 238 |
| 8. Perszeveráció..... | 238 |
| 9. Anticipáció..... | 239 |
| 10. Metatézis..... | 241 |
| 10. a) Az egyik felcserélt elem nem jelenik meg..... | 241 |
| 11. Egyszerű nyelvbotlás..... | 242 |
| 11. a) Betoldás..... | 242 |
| 11. b) Csere..... | 242 |
| 11. c) Kiesés..... | 242 |
| 12. Többféle osztályozás..... | 243 |
| A pedagógusjelöltek Kazinczy-versenye | 248 |
| A pedagógusjelöltek Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 34. országos döntője • <i>S. Novotny Júlia – Pelczéder Katalin</i> | 248 |
| Simonyi Zsigmond: A nyelvtudományról (A verseny kötelező szövege) 250 | |
| A 2006-os döntő kötelező szövegének értelmezése és előadásának egyfajta megközelítése • <i>N. Markó Julianna</i> | 251 |
| A magyar és a nemzetközi tudományos élet eseményei..... | 254 |
| 11. Nemzetközi Klinikai Fonetikai és Nyelvészeti Konferencia • <i>Gósy Mária</i> | 254 |
| 5. nemzetközi konferencia a beszéd motoros kontrolljáról • <i>Gósy Mária</i> | 255 |
| Az Amerikai Akusztikai Társaság 151. konferenciája • <i>Bóhm Tamás</i> | 257 |
| Beszéd kutatás 2007 tudományos ülészak (Felhívás)..... | 259 |
| A kötet szerzői | 260 |

95 ÉVES LENNE VÉRTES O. ANDRÁS (1911–1997)



A felvétel 1995. december 4-én, a „Kruzsók”-ban készült
(a Gellért sörözőben)

Szép kort ért meg – ahogy mondani szokás; 86 évesen hunyt el egy örömben gazdag és bánatokban is bővelkedő élet után. Máshogyan fogalmazva: emberi életet élt, mégpedig példás és példamutató emberi életet. Aki ismerhette, tudja, hogy ezek a szavak nem a „formai előírás”, avagy az emlékezet „hiányosságai” következtében íródtak. Vértess O. András mint ember, mint kutató, mint pedagógus, mint kolléga, mint atyai jó barát sokaknak mutatott példát szakmai és emberi tekintetben egyaránt.

Közhelyszerűen rohanó világunkban azonban nem feltételezhetjük, hogy a miénknél fiatalabb generációk is biztosan ismerik a nevét. Bár szeretném, ha így lenne. Ki volt ő?

Talán a legelégedettebb akkor lenne, ha azt írnám, hogy pedagógus, mint az édesapja. Hatvanhat évesen életéről beszélve megemlíti, hogy eleinte a földrajz és a matematika érdekelt, de már 13–14 évesen a nyelvészet iránt kötelezte el magát. (Mindez a saját szavaival is hallható honlapunkon, a *Fonetika Interneten*: <http://fonetika.nyud.hu/nyelveszek2005/vertesoandras.htm>.) Az érettségét (Werbőczy István Reálgymnázium, Budapest, Krisztinaváros, 1929; az iskola 1946 óta a Petőfi Sándor Gymnázium nevet viseli) követően tanulmányait a budapesti Pázmány Péter Tudományegyetem Bölcsészettudományi Karán végezte, magyar–francia szakon. Tudósi pályája már az egyetemi évek alatt elkezdődött;

Gombocz Zoltán tanítványaként az egyetemi Magyar Nyelvtudományi Intézet szeniora volt (1933–1935), ez a „beosztás” akkor tanársegédi feladatokat jelentett. Az egykori Gombocz-szeminárium ablaka a Múzeum körút fájra nézett; ma a Bölcsészettudományi Kar nyelvészeti tanszékei ismét ugyanezen a helyen vannak. Elsőévesen jelent meg első tanulmánya, majd igen korán hangtörténettel kezdett foglalkozni. Melich Jánosnál doktorált. Később a nyelvtudomány kandidátusa (CSc), majd a nyelvtudomány doktora (DSc) címet is megszerezte. Első munkahelye az újpesti szövő-fonóipari szakiskola, ahol magyart tanított (1936–1948). Visszaemlékezésében azt mondja, hogy akkor legjelentősebb feladatának az osztályfőnöki munkát tekintette, vagyis a pedagógiai munkát, a gyermekek nevelését. Közben (1946-tól) a budapesti műszaki középiskolák, felső ipariskolák magyar nyelvi és irodalmi szakfelügyelője. 1948-tól a Gyógypedagógiai Tanárképző Főiskola oktatója, ahol fonetikát, anyanyelv-elsajátítást és más tárgyakat adott elő. 1957-ben elkerült a főiskoláról (ennek dokumentumát lásd *Magyar Fonetikai Füzetek* 23. 1991: 14). Ezt követően az MTA Nyelvtudományi Intézetében dolgozott tudományos munkatársként nyugdíjba vonulásáig (1976). Számos értékezésnek volt értő bírálója; tőle a kritikai megjegyzések is úgy hatottak, mint mástól a dicsérő szavak. Véleményével mindig segíteni akart, utat mutatott, avagy kérdéseket vetett föl. A gyógypedagógiai főiskolával élete végéig megmaradt a kapcsolata, a fonetika és a „gyógyító pedagógia” iránti érdeklődése motiválta mindvégig a munkásságát. Ennek megfelelően kutatási területei elsősorban a fonetika, főként a történeti vonatkozások és a gyógyító pedagógia voltak; de érdekelték a nyelvleléktan, a gyermeknyelvnek, a beszéd változásainak, a hangstilisztikának és a beszédhigiénének a kérdései is. Nem egy közös munkálatban vett részt az intézetben töltött évek alatt; ilyen volt a magyar nyelv finn-ugor elemeinek szótára, avagy a leíró magyar hangtan fejezetei. Számos kérdést elsőként tárgyalt a magyar fonetikában, például az időtényezők stílusbeli szerepét vagy a hümmögés (nála: *hmgetés*) problémáját.

Különböző tanulmányait, *A magyar leíró hangtan története az újgrammatikusokig* (Akadémiai Kiadó, Budapest, 1980) című könyvét, hangstilisztikai kötetét (*Bevezetés a magyar hangstilisztikába*. NytudÉrt. 124. Akadémiai Kiadó, Budapest, 1987) a mai napig rendszeresen idézik. Úgy gondolom, rajta kívül senki nem tudta volna a magyar fonetika történetét olyan színvonalon megírni, ahogyan ő tette. Ahhoz ugyanis nemcsak a latin, a német és más nyelvek ismerete, a fonetika tudományának elkötelezett és hozzáértő művelése volt szükséges, hanem végtelen türelem, aprómunka, szakértelem és alázat a tények iránt; csupa olyan tulajdonság és tudás, amely csak kivételes embereknek adatik meg.

Vértes O. András a Magyar Fonetikai, Foniátriai és Logopédiai Társaság egyik alapítója, 1981 és 1985 között elnöke, 1995-től a Társaság elnökségének örökös tagja. A Finn Tudományos Akadémia akadémikusai sorába választotta.

Az MTA Nyelvtudományi Intézetének Fonetikai Osztálya (Kempelen Farkas Beszédkutató Laboratórium) 1997-ben egynapos konferenciát szervezett Vértes O. András tiszteletére a Magyar Tudományos Akadémián. Erről, a szándékról és

a programról még tudott Vértes O. András, azt azonban nem gondoltuk volna, hogy mire az eseményre sor kerül, „András bácsi” már nem lesz velünk. Néhány hónappal korábban ment el. A tiszteletére tervezett konferencia már csak az emlékének szólhatott...

Az akkor elmondott bevezető szavaimban idéztem egy kis történetet a gyermek Vértes O. András „első diagnózisáról”, amit húga, Vértes O. Augusztina mesélt el nekem. Itt most azért kap helyet, hogy megőrződjön. Édesapjuk, Vértes O. József (aki az ideges gyermekek állami középiskolájának volt a vezetője) egy kisfiút és az apját várta egy alkalommal, hogy beszélgessenek a gyermekről. Megkérte az akkor 5 éves Andrást, hogy míg ő az apával beszél, játsszon a kisfiúval, aki – mint mondta – egy „ideges gyermek”. A kis Vértes O. András eleget is tett egy ideig édesapja kérésének, majd amikor a gyermek már többször hozzávágta a játékokat, beszaladt a szobába, és közölte: „Ez nem ideges, hanem bolond!”

A *Magyar Fonetikai Füzetek* 23. kötete a *Studia in honorem Andreae O. Vértes* címet viseli, a 80. születésnapra készült.

A *Beszéd kutatás* – a fonetikával, beszédpatológiával és beszédtechnológiával foglalkozó szakemberek évente megjelenő gyűjteményes kötete – ez évi összeállítását tisztelegés Vértes O. András emléke előtt. Az ő tanítványaként igyekszem a beszéddel kapcsolatos kutatásokat úgy folytatni és irányítani, ahogyan arra tőle közvetve és közvetlenül is indítást kaptam. A szellem folyamatosan köztünk van, a gondolatai pedig még nagyon sokáig fognak termékenyítőleg hatni a beszéddel foglalkozó kutatókra.

Vértes O. András sokakat mondhat tanítványának, de még többen vannak, akik tanítómesterüknek vallják őt. Nem csupán szigorúan szakmai értelemben, de az emberi tisztességet, az erkölcsöt és a tudományos etikát tekintve is.

Gósy Mária

A SEMLEGES MAGÁNHANGZÓ NYELVI FUNKCIÓI

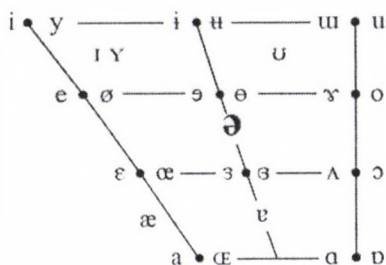
Gósy Mária

Bevezetés

A hangszalagokat megrezegtető levegő eredménye a zönge vagy elsődleges (primér) hang, amely a toldalékcsofen módosul az ott lévő beszédszervek nyelvspecifikus működése következtében. A szájüreg alakja változik attól a parancstól függően, amit az agy küld egy adott magánhangzó-minőség létrehozása céljából. Akár artikulációs, akár akusztikai vagy percepciósi szempontból tekintjük, a magánhangzók fonetikai meghatározása nehezebb, mint a mássalhangzóké, mivel nehéz az adott képzési konfigurációjuk pontos leírása. Ez pedig következményesen jelentkezik az akusztikumban, sőt a percepcióban is. A magánhangzók képzésére a teljes szájüreg a rendelkezésünkre áll, a biológiai rendszer relatív rugalmassága nagymértékben érvényesülhet az artikuláció folyamán. A beszédképzés stabilitását magyarázó kvantális elmélet szerint (Stevens 1972) egy nyelv artikulációs mozgásai között léteznek olyan igen kicsi változások, amelyeket az ember nem észlel. Az artikuláció tehát rugalmasabban mehet végbe, hiszen a nyelvspecifikus észlelési folyamatok biztosítják a nyelvi azonosságot. Az ejtés nem észlelhető „zónái” a kvanták. A kvantális elmélet nem tér ki azonban az egyéni ejtési sajátosságokra, amelyek még egy nyelven belül is tovább növelhetik a magánhangzó artikulációjának pontos meghatározását.

A magánhangzók leírásának problémáját a neves angol fonetikus, Daniel Jones úgy igyekezett megoldani, hogy megalkotta a kardinális magánhangzók rendszerét (1922). Ebben a rendszerben meghatározott észlelési referenciapontok szolgálnak az egyes magánhangzók képzési módjának elvi definiálására. Jones rendszerében eredetileg nyolc magánhangzó található, amelyek nem egy adott nyelv magánhangzói, hanem a szájüregi lehetőségek és az észlelés összefüggésének eredményeképpen meghatározott magánhangzó-lehetőségek. A pontosabb, részletesebb leírás érdekében Jones létrehozta a másodlagos kardinális magánhangzók rendszerét további 22 típussal (vö. Gósy 2004). A kardinális magánhangzók fő kritikája azon alapszik, hogy összekeverednek az artikulációs és az észlelési kritériumok, de további vitatható pontjai is vannak. Mindennek ellenére a mai napig ez a legjobban és legelterjedtebben használható rendszer egy nyelv magánhangzóinak leírására. Ennek a magánhangzó-rendszernek a közepén található a **semleges magánhangzó**, a svá, amely középponti helyet foglal el (a szájüregi térben) abban az értelemben, hogy ejtésekör a nyelv a szájüregben – mindkét irányban – lényegében közepén van. Szabatosabb megfogalmazásban az mondható, hogy az [ə] magánhangzó képzésekör a vokális csatornára a sem-

leges konfiguráció a jellemző, amelynek megfelelően az ajkak nem kerekítettek (1. ábra).



1. ábra

A kardinális magánhangzók és a semleges magánhangzó (svá) helyzete

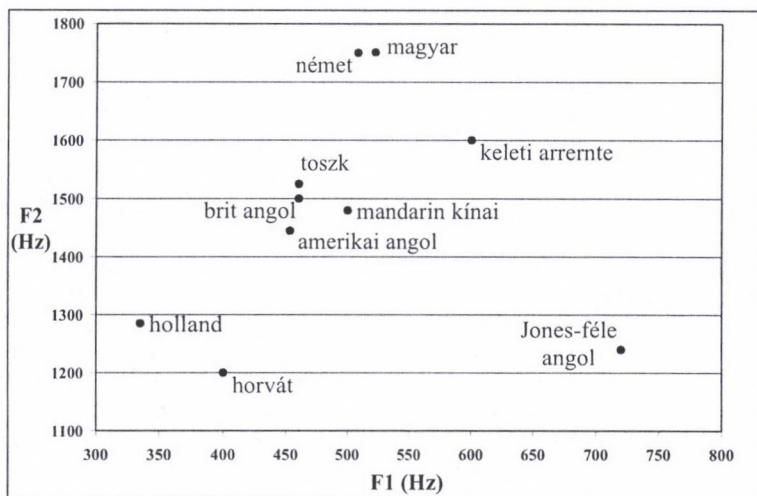
A semleges magánhangzót a szakirodalomban sokféleképpen nevezik, *homályos*, *meghatározhatatlan*, *redukált*, *centrális*, *mormolt*, illetve *hangsúlytalan szótagi magánhangzó*-nak (az angol szakirodalomban: *obscure vowel*, *indeterminate vowel*, *reduced vowel*, *central vowel*, *medial tongue vowel*, *murmured vowel*, *unstressed syllable vowel*). A leggyakoribb megnevezése hagyományosan mégis a *svá*.

A *svá* szó maga a héberből ered, a jelentése ‘semmisség’. Az angolban elsőként Peter Giles használta egy filológiáról szóló tankönyvében 1895-ben (*A Short Manual of Comparative Philology for Classical Students*). A megfordított *e* betű ugyanakkor sokkal régebben megtalálható a fonetikában a mai fonetikai szimbólumértékkel. Így elsőként Johann Andreas Schmeller használta 1821-ben. A *svá* szó azonban Giles tankönyvéig nem fordult elő (MacMahon 2003).

A magyar hangtan történetében hárman foglalkoztak a „határozatlan hangszínnű” magánhangzókkal. Kyss Sándor *vocalis elementaris*-nak nevezte, és úgy vélte, hogy orrhangú magánhangzó, amelyre a – bár nem kizárólagos – zárt szájjal való ejtés a jellemző. Szerinte csupán négy mássalhangzóval alkothat szótagot, ezek a [l, r, m, n]. Matúsik András pontosabban írja le a semleges magánhangzót. Nem tekinti orrhangúnak, s a zárt ejtését sem tartja annyira jellemzőnek. Mindketten utalnak arra, hogy a nyögő betegek is ezt a hangot hallatják, valamint hogy a sírásban, a nevetésben, a morgásban is ez a hang hallatszik. Jellegetes beszédbeli szerepe az igenlésben, a tagadásban, a csodálkozás kifejezésében érhető tetten (Vértes 1980: 119). Kerekes Ferentz fontos megállapítása, hogy a felpattanó zárhangok ejtésével hozza kapcsolatba a semleges magánhangzó megjelenését, továbbá megemlíti a héber „scheva” hangot, amelyhez hasonlónak mondja (1836). Érdekesség, hogy a 19. század elején még az is vita tárgya volt, hogy az [ø, y] magánhangzók összetett hangok, azaz kettőshangzók vagy nem (Vértes 1980: 106). Ugyanakkor a 20. századi nyelvjárási gyűjtemések kétféle [ø] magánhangzót azonosítottak; a legfőbb különbség köztük a „szük-

nek”, illetve „szélesnek” nevezett ejtés. Feltételezték, hogy ezek korai nyelvjárási fonémák nyomait őrzik. Felmerül itt is esetleg egy svászerű ejtés lehetősége. A semleges magánhangzó említése a magyar fonetikákban a 20. század végéig csak elvétve fordul elő.

A semleges magánhangzó artikulációs konfigurációja látszólag egyértelműen meghatározott, mégis más magánhangzókhoz képest sokkal nagyobb a megvalósításának a variabilitása. A nyelv vízszintes és függőleges helyzete különféle viszonylatokban (pl. kissé magasabban és hátrább vagy éppen előrébb a szájüregben) is eredményezheti a semleges magánhangzó észleletét. A világ nyelveiben meglehetősen nagy eltérésekkel jellemezhetők a svák. Az artikulációs sokféleségnek a következményei az első és a második formánsok jelentős eltérései (2. ábra).



2. ábra

A svá első két formánsának értékei különböző nyelvekben
(A brit angolra Jones adatait is feltüntettük.)

Az a tény, hogy a semleges magánhangzóra közölt akusztikai adatok meglehetősen eltérőek, két alapvető kérdést vet fel. Meghatározható-e egyáltalán a semleges artikuláció, létezik-e ilyen konfiguráció? Feltételezhető-e, hogy a semleges magánhangzó képzése nyelvspecifikus? Ez utóbbi esetben a „semleges artikuláció” mint olyan csupán elméletileg létezne, valójában mindig az adott nyelv magánhangzó-rendszere, illetőleg a nyelvet beszélők észlelése határozná meg a semlegesnek tekintett állapotot.

A svának azonban csak egyik problémája az artikuláció. Ez a magánhangzó a szokásosnál több nyelvi funkcióval rendelkezik egy nyelven belül és több nyelv

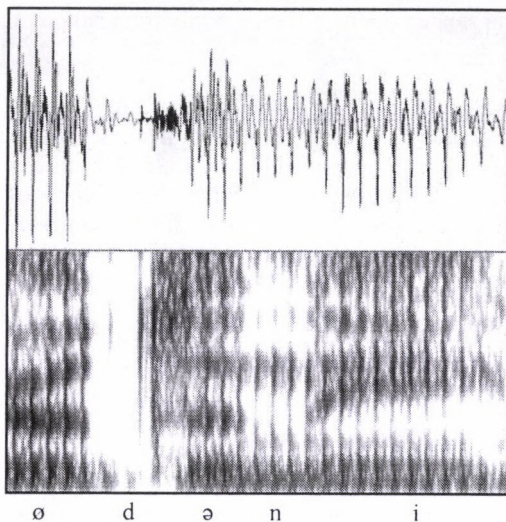
tekintetében is. Ha a nyelvi funkciót akarjuk meghatározni, azzal szembesülünk, hogy minden más, a verbális kommunikációban használt jelnél több funkciót kell a svához rendelnünk. A szakirodalom a semleges magánhangzó számos funkcióját nevezi meg a különböző nyelvekben, mégsem tűnik teljesnek a felsorolás. A következőkben megkíséreljük áttekinteni a semleges svá artikulációs, illetőleg akusztikai sajátosságait, valamint nyelvi funkcióit a magyarban.

A svá funkciói

A különféle nyelvek magánhangzó-rendszerét áttekintve, azt látjuk, hogy a [ə] lehet fonéma és allofón is (Jones 1966, Koopmans-van Beinum 1994, Laver 1994, Barry 1995, Ladefoged–Maddieson 1996, Hirschfeld–Wallraff 2002, Granser–Moosmüller 2003). Az IPA-kézikönyvben (Handbook 1999) található nyelvek 42%-ában a semleges magánhangzó fonéma alapváltozata például az amharik, a sindi vagy a szlovén nyelvben, avagy mellékvariáns, mint például a horvátban, a katalánban, a héberben és az írben. A németben egyidejűleg fő- és mellékvariáns is. A héber sajátos alkalmazását mutatja e magánhangzónak. A svá ekkor a magánhangzó hiányát jelzi (MacMahon 2003). A sváepentézis például egy absztrakt fonológiai folyamat eredménye a hollandban (Warner et al. 2002). Az angol spontán beszédben 22,9%-ban, a németben mintegy 30%-ban, a franciában 7,6%-ban jelenik meg a semleges magánhangzó más magánhangzók realizációjaként (Onishi 1981). Számos nyelvben, mint például a spanyolban és a magyarban is, a semleges magánhangzó nem fonéma. Jellemzően előfordul ugyanakkor a spontán beszédben mint más magánhangzók helyettesítője (Harmegnies–Poch-Olivé 1992; Gósy 2004). Ennek oka egyfelől a beszédtempóban, másfelől az elnagyolt artikulációban keresendő. A semleges magánhangzónak eddig tehát a következő funkcióit láttuk különféle nyelvekben: fonéma (alapváltozat és mellékvariáns), a magánhangzó hiányának jelzője, fonológiai folyamat eredménye, más magánhangzó realizációja, azaz helyettesítője a spontán beszédben. A továbbiakban azt elemezzük, hogy a magyarban milyen funkciók jellemzik a semleges magánhangzót.

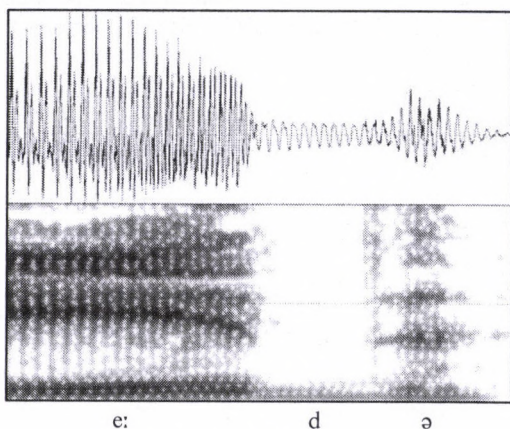
A magyar svá más magánhangzó megvalósulása is lehet a spontán beszédben, ami azt jelenti, hogy a kívánt magánhangzó-minőség helyett a beszélő a svát ejti. Ezt helyettesítő funkciónak tekintjük. Megjelenik koartikulációs következményként; ekkor a fiziológiai kényszer eredménye, vagyis a beszélő szándéka ellenére ejti bizonyos hangkörnyezetben a semleges magánhangzót. Ez utóbbinak három jellegzetes kontextusfüggő előfordulása van (sporadikusan egyéb hangkapcsolatokban is előfordulhat): a) a felpattanó zárhang és a nazális kapcsolatában (pl. *gnú* = g[ə]nú vagy *teknő* = tek[ə]nő); a 3. ábra a *köpköd[ə]ni* szó hangszínképéből az -*ödni* részt mutatja. b) Bizonyos mássalhangzók után, abszolút szóvégi helyzetben (pl. [b, d, g, m, n]# → *adag[ə]*); a 4. ábra a *véd[ə]* szó ejtéséről készült spektrogram. Végül c) a pergőhang képzésében (vö. az 5. ábra *róka* és a 6. ábra *fűrész* szavának regisztrátumait).

Mindegyik – az ábrákon szemléltetett – előfordulásban a semleges magánhangzó mintegy ejtéskönnyítőként, illetve az ejtés lehetséges vagy szükséges megvalósításaként jön létre. Az ábrák hangszínképei és rezgésképei egyértelműen igazolják a semleges magánhangzó jelenlétét.



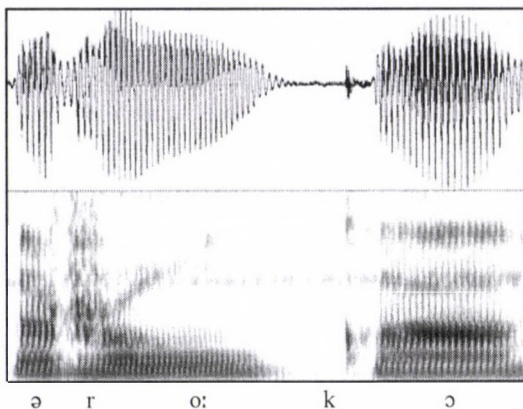
3. ábra

A semleges magánhangzó a *köpköd[ə]ni* szóban
Az ábra az -ödni részt szemlélteti.



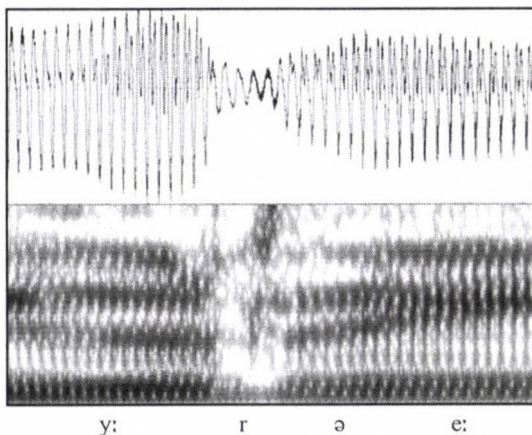
4. ábra

A semleges magánhangzó a *(v)éd* szóban mondatvégi helyzetben



5. ábra

A svá megjelenése a pergőhangban abszolút szó eleji helyzetben (*róka*)

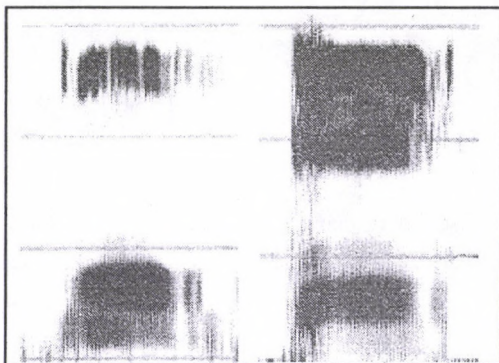


6. ábra

A svá megjelenése a pergőhangban intervokális helyzetben
(*-űr- a fűrész szóban*)

Számos esetben a beszélő értelmi, érzelmi vagy egyéb okból hangsúlyozni kíván valamit a beszéd során. Ilyenkor szinte kivétel nélkül megjelenik a semleges magánhangzó mintegy fiziológiai kényszerként. Például egy haragosabb, mérgesebb ejtésben az *Add ide!* közlés első szavának végén hallható is a semleges magánhangzó artikulációjának következménye. Hasonló következménnyel jár a szóhatárjelzés bizonyos szegmentális homonímia kialakulása esetén, például *arab lány* de *a rab lány* vagy *halálok* de *halál ok*. Egy egészen más funkcióban

találjuk a semleges magánhangzót az anyanyelv-elsajátítás gögicselési szakaszában. Ekkor protomagánhangzó, vagyis olyan „nem mássalhangzó”, amely gyakorlatilag bármely más magánhangzó-minőség kiváltására képes (7. ábra). E protomagánhangzós funkciójú svá lényegében úgy jön létre, hogy a csecsemő zöngét képez, a nyelve azonban semleges helyzetben, mintegy középen van a szájüregben, s az ajkai kismértékben nyílnak szét. A felnőtt környezet pedig a hangélményt a semleges magánhangzóval azonosítja.

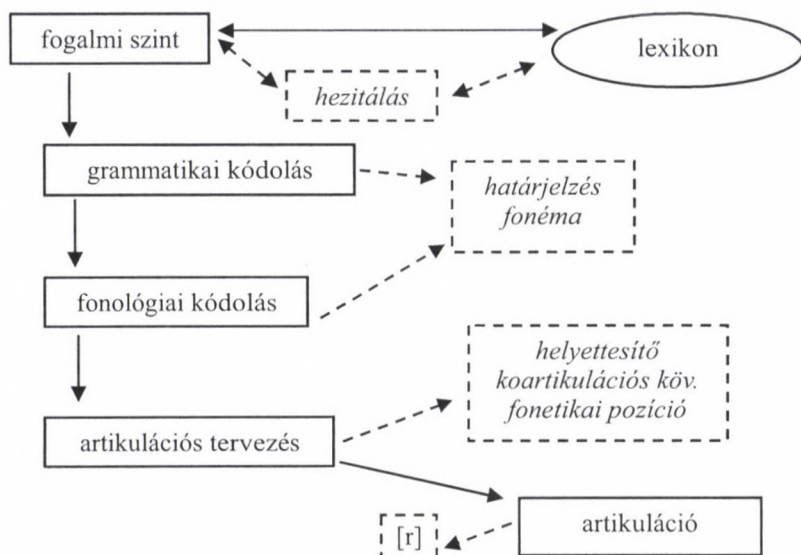


7. ábra

Gögicsélt, semlegesnek azonosított magánhangzó (a bal oldalon) és [ø]-nek azonosított, gögicsélt magánhangzó (a jobb oldalon) hangszínképe azonos életkori szakaszban (3000 Hz-es tartományban)

A spontán beszéd egyik jellemzője, hogy a beszélő a tervezési bizonytalanság, illetőleg az átmeneti szótalálási nehézségek következtében megakad, s ez nem egyszer hangadással jár együtt. Ezt nevezzük hezitálásnak. A hezitálás leggyakrabban svászerű hanggal történik, vagyis a semleges magánhangzó egy újabb funkcióban, a beszédtervezési folyamat problémajelzőjeként jelenik meg. A magyar anyanyelvűek hezitálására (vö. *őzés*) különösen jellemző a svá képzése. A magyarban összesen tehát nyolc különböző funkcióban fordulhat elő a semleges magánhangzó. A funkcionális különbözőségek a beszédprodukciós folyamat különböző szintjeihez kapcsolhatók (8. ábra).

A fonéma (ez a magyarra nem jellemző), valamint a szóhatárjelzés funkcióban a semleges magánhangzó ejtése a grammatikai és a fonológiai kódolással függ össze. A koartikulációs következmények és a más magánhangzók realizációi, a helyettesítések a fonológiai és az artikulációs tervezéshez köthetők. A szóvégi helyzetben megjelenő svák az artikulációs tervezéssel vannak kapcsolatban, míg a pergőhang részeként az aktuális artikuláció eredményei. Ez utóbbi esetben a hangszalagok már rezegnek, mielőtt a [r] mássalhangzóra jellemző nyelvhegyi pergetések elkezdnének, s ebben a szakaszban a nyelv mintegy semleges helyzetben van a szájüregben.



8. ábra

A semleges magánhangzó megjelenése a beszédproduktions folyamatban
(funkcionális szemléletben)

Noha a svá magánhangzó felsorolt funkciói vitathatatlanul léteznek a beszédben, fonetikai szempontból választ kell adnunk arra a kérdésre, hogy a funkcióknak megfelelően van-e különbség az egyes realizációk képzésében. Tekintve, hogy az akusztikai szerkezetből visszakövetkeztethetünk az artikulációra, akusztikai-fonetikai kísérletek sorozataiban elemeztük a semleges magánhangzót.

Az ábráról leolvasható, hogy a svá megjelenése a beszédproduktions folyamat egyes szintjeihez (pl. artikuláció), avagy több szinthez is hozzárendelhető (pl. a hezitálás). Nem tüntettük fel a gögicsélt svákat, amelyek egyfelől ugyan az artikuláció következményei, másfelől azonban mint szándékolt protohangok bizonyos fokig a kognitív szinthez is kapcsolhatók.

Kísérleti személyek, anyag, módszer

Összesen 10 fiatal felnőtt (21 és 30 év közöttiek) és egy fiúsecsemő (1 és 10 hónapos kora között) vett részt a kísérletben. A beszédanyag 120 izolált szót és szókapcsolatot, egy 25 perces spontánbeszéd-anyagot és a csecsemő 3 órányi gögicselését tartalmazta. Az izolált szavak összeállításakor ügyeltünk arra, hogy a fiziológiai kényszer következtében elvárható fonetikai helyzeteket tartalmazák, illetőleg hogy a vizsgálandó koartikulációs jelenségek előforduljanak. Az anyagban a határjelző funkcióra alkalmas helyzetek is voltak. Az akusztikai fo-

netikai elemzéseket a CSL 4300B digitális jelfeldolgozóval, valamint a Praat programmal (4.2) folytattuk. A statisztikai elemzéseket az SPSS 8.0 verziószámú változatával készítettük (ANOVA, párosított *t*-próba) 95%-os szinten.

Eredmények

A semleges artikulációnak megfelelő akusztikai szerkezet szerint modellszinten az első formáns értéke 500 Hz, a második formánsé 1500 Hz, a harmadiké pedig 2500 Hz (Pickett 1980). A természetes ejtésben ezek az értékek valamenynyire módosulnak, hiszen biológiai rendszerről van szó. Ekkor az első formánsra kapott átlagérték 475 Hz, a második formánsé 1450 Hz, a harmadiké 2375 Hz. Egy másik hipotetikus modellt, amelyik a fonológiai jegy és a formánsértékek közötti összefüggéseket elemzi, a következő értékeket adja a svára: $F_1 = 370$ Hz, $F_2 = 1630$ Hz, $F_3 = 2600$ Hz (ten Bosch 1991: 45). A szakirodalomban két lényeges pontot hangsúlyoznak a semleges magánhangzó akusztikai szerkezetére vonatkozóan. Az egyik az, hogy a svá akusztikai variabilitása a második formánsban érhető tetten. A másik pedig az, hogy a fonémaértékű svá magánhangzó F_2 -jének értékei kismértékben periférikusnak tekinthetők. A keleti arnerntében (Ausztráliában beszélt nyelv) például ez az érték 1600 Hz körüli, míg a pekingi mandarin kinaiban 1220 Hz táján realizálódik (Granser–Moosmüller 2003).

Magyar spontán beszédben elemzett semleges magánhangzók formánsértékeit a középső nyelvallású, elől képzett, ajakkerekítéses magánhangzókéval együtt adja meg az 1. táblázat. A táblázat az F_1 és az F_2 értékeit mutatja. A kétféle elemzett magánhangzó első és harmadik formánsainak frekvenciaértékeit összevetve, nem látunk szignifikáns, matematikailag igazolható különbséget. Az [ø] és az [ə] második formánsai azonban statisztikailag különböznek egymástól; ami azt jelenti, hogy az eltérő magánhangzó-minőség az F_2 értékében ragadható meg ($t(119) = -2,072$, $p < 0,042$).

1. táblázat: Az [ə] és [ø] magánhangzók magyar spontán beszédben

| Magánhangzók | F_1 (Hz) | | F_2 (Hz) | |
|--------------|------------|-----------------|------------|-----------------|
| | Átlag | Átlagos eltérés | Átlag | Átlagos eltérés |
| [ə] | 549,07 | 109,52 | 1787,50 | 149,31 |
| [ø] | 460,12 | 75,10 | 1605,94 | 185,91 |

Az elemzett spontán beszédben a svák mintegy 40%-ban jelennek meg egyéb magánhangzók helyett, függően az adott artikulációtól, illetőleg a beszélő beszédtempójától. Hasonló arányt találtak holland spontán beszédben is (van Bergem 1995). A semleges magánhangzó – várhatóan – főként hangsúlytalan helyzetben jelenik meg, azonban az előfordulások mintegy negyede a hangsúlyos magánhangzókat is érinti (Gósy 2004).

A 2. táblázatban a spontán szöveg eredeti magánhangzóinak és a helyettesítő funkciójú sváknak a formánsértékei láthatók. A formánsok értékei között szigni-

fikáns különbséget találtunk ($t(59) = 3,208, p < 0,002$), azonban az időtartam tekintetében ez nem állt fenn.

2. táblázat: Négy magánhangzó időtartamai és formánsértékei eredeti ejtésben és a svával helyettesített realizációkban spontán beszédben

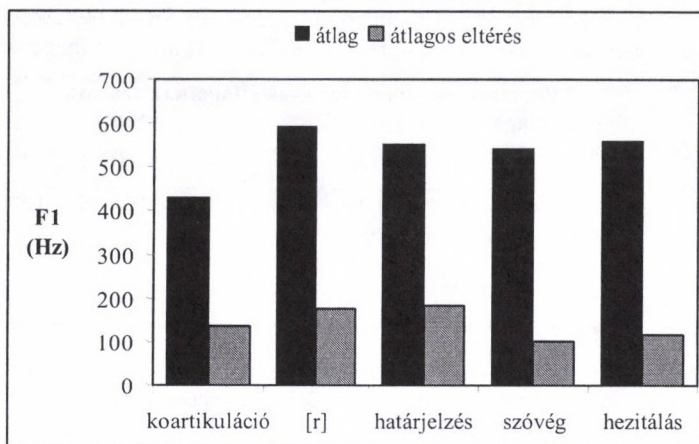
| V | Időtartam- és formánsértékek átlagainak szórása | | | | | |
|-----|---|---------------------|---------------------|----------------|---------------------|---------------------|
| | Eredeti magánhangzó | | | Svá | | |
| | Időtartam (ms) | F ₁ (Hz) | F ₂ (Hz) | Időtartam (ms) | F ₁ (Hz) | F ₂ (Hz) |
| [o] | 30–90 | 340–510 | 720–1100 | 40–85 | 310–500 | 1075–1600 |
| [ɔ] | 50–210 | 480–600 | 960–1230 | 25–155 | 390–540 | 1160–1500 |
| [a] | 80–155 | 600–760 | 1330–1510 | 50–124 | 395–620 | 1415–1770 |
| [ɛ] | 47–200 | 480–585 | 1460–1835 | 28–130 | 370–570 | 1060–1785 |

Az [o] időtartamának átlaga 65,4 ms, a helyettesítői 58,3 ms; az F₁/F₂ átlagértéke 413,7/920,3 Hz, helyettesítésben 442,3/1453,6 Hz. Az alsó nyelvvállású magánhangzóé 158,7 ms, a helyettesítésben 149,2 ms; az első/második formáns átlaga 560,6/1180,2 Hz, helyettesítésben 510,4/1415,9 Hz. A legelső nyelvvállású magánhangzóé 122,8 ms, a helyettesítőé 118,5 ms; az F₁/F₂ átlagértéke 712,9/1460,6 Hz, helyettesítésben 588,2/1683,9 Hz. A palatális magánhangzóé 129,3 ms, a helyettesítőé pedig 105,4 ms; az első/második formáns átlaga 554,4/1732,8 Hz, helyettesítésben 510,7/1488,5 Hz.

A részletekbe menő elemzésekkel arra kívántunk választ kapni, hogy vajon a különböző nyelvi funkciók különféle artikulációval valósulnak-e meg. Változik-e tehát a svák időtartama és/vagy formánsszerkezete az adott nyelvi funkció függvényében? A semleges magánhangzó temporális jellemzőinek adatai azt igazolták, hogy szignifikáns a különbség a hezitáláskor előforduló és a más nyelvi funkciójú svák között ($t(69) = -3,708, p < 0,003$). A hezitált svák átlagos időtartama 675,88 ms volt (átlagos eltérés 443,52 ms), míg szóhatárjelző funkcióban az átlag 80,88 ms (átlagos eltérés 20,46 ms), a koartikuláció esetén 38,74 ms az átlag (az átlagos eltérés 9,30 ms), az abszolút szóvégi helyzetre pedig átlagosan 71,92 ms adódott (az átlagos eltérés 25,15 ms). A [r] mássalhangzó részeként a svák időtartamát átlaga 52,76 ms (az átlagos eltérés 15,80 ms). A semleges magánhangzók időtartama ugyancsak szignifikáns eltérést mutat a fentebbi helyzetek és azon előfordulások között, amikor más magánhangzót helyettesítenek a spontán beszédben ($t(149) = -3,470, p < 0,001$).

A formánsszerkezet elemzése is azt mutatja, hogy a sváartikuláció funkciófüggő. Várhatóan nincs azonban annyiféle artikulációs konfiguráció, ahányféle funkcióban a semleges magánhangzó megjelenik. Az első formánsok mérési adatai azt mutatták, hogy a koartikuláció eredményeként megjelenő svák frekvenciaértéke különbözik jelentősen az összes többitől (9. ábra). Az előbbieket át-

lagértéke 433,7 Hz (az átlagos eltérés 44,3 Hz), míg az összes többié 565,2 Hz (az átlagos eltérés 69,9 Hz).



9. ábra

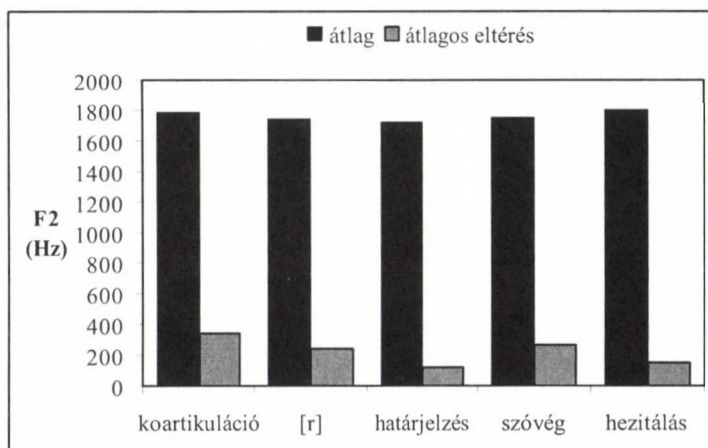
A semleges magánhangzó első formánsai (átlag, átlagos eltérés) különböző funkciókban

Az F_1 szórása legnagyobb a határjelzés esetén és a pergőhang részeként (ez utóbbinál az egyéni különbségek indokolják a tartomány megnövekedését). A legkevesebb artikulációs módosulást a nyelv függőleges mozgásában – az akusztikai adatokból következtetve – a szóvég esetén tapasztaltuk, ami ennek a helyzetnek a fonetikai sajátosságaival jól magyarázható.

Az F_2 adatai nem mutatnak lényeges eltéréseket az átlagokban, a szórástartomány azonban különböző (10. ábra). Ez az eltérés valószínűsíthetően a funkcionális különbségek hordozója. A legszűkebb a határjelzés és a hezitálás esetében, a legszélesebb pedig a koartikulációban; a különbségek azonban nem szignifikánsak. Matematikailag igazolható ugyanakkor a második formánsok különbsége a koartikuláció, a hezitálás és a helyettesítő funkció esetében ($F(2, 72) = 23,471; p < 0,0001$). Hasonló eredményeket kaptak a második formánsokra vonatkozóan a mandarin svá magánhangzók elemzésekor különböző kontextusokban (szótag végi helyzetben, koartikulációban az alveoláris és a veláris nazális mássalhangzókkal, vö. Sproat 1998).

A harmadik formánsok elemzése nem mutatott lényeges eltérést a funkciótól függően. Nem is vártunk itt jelentős különbségeket. A konkrét értékek nagyon hasonlóak az egyes funkcióban vizsgált svák esetében. A [r] ejtésékor a semleges magánhangzó harmadik formánsának átlaga 2859,6 Hz (az átlagos eltérés 246,8 Hz); a határjelzés esetén az átlag 3005,7 Hz (az átlagos eltérés 121,2 Hz); a koartikulációban az átlag 2887,6 Hz (az átlagos eltérés 342,2 Hz); szóvég

helyzetben az átlag 2981,3 Hz (az átlagos eltérés 265,7 Hz); míg a hezitálásokban az átlagérték 2884,9 Hz (az átlagos eltérés 147,3 Hz).



10. ábra

A semleges magánhangzó második formánsai (átlag, átlagos eltérés) különböző funkciókban

Ha mind az első, mind a második formáns értékeit egyidejűleg vesszük tekintetbe, akkor csupán a koartikuláció és a hezitálás, valamint a szóvég és a határjelzés semleges magánhangzóit között találunk matematikailag is igazolható akusztikai különbséget. Ez azt jelenti, hogy a beszélők talán nem tudatosan ugyan, de bizonyos fokig eltérően ejtik a svát némely nyelvi funkció esetén. Feltételeztük, hogy a hezitálásokban és a határjelzésekben megjelenő semleges magánhangzók különbözők lesznek. A hezitálás ugyanis mint megakadásjelenség jelentősen elkülönül a többi szegmentális szintű svá magánhangzótól; a határjelzés esetén pedig bizonyos szupraszegmentális sajátosságok is megjelen(het)nek, például a hangsúly, s ez jelentősebben módosít(hat)ja a svá képzését. Nem feltételeztük ugyanakkor, hogy a szóvégi helyzet is jelentős artikulációs módosulást okoz.

Feltételeztük, hogy a semleges magánhangzó a koartikuláció esetén különböző artikulációs konfigurációt mutat azoktól a sváktól, amelyek a hezitálásokban, illetőleg más magánhangzók helyettesítőiként jelennek meg. Az első és a harmadik formánsok értékei nem mutattak jelentős eltérést, azonban a második formánsok szignifikánsan különböztek. Ez a nyelv vízszintes mozgásának eltéréseire utal (a Kruskal–Wallis-próba szerint az eltérés majdnem szignifikáns, $p < 0,055$).

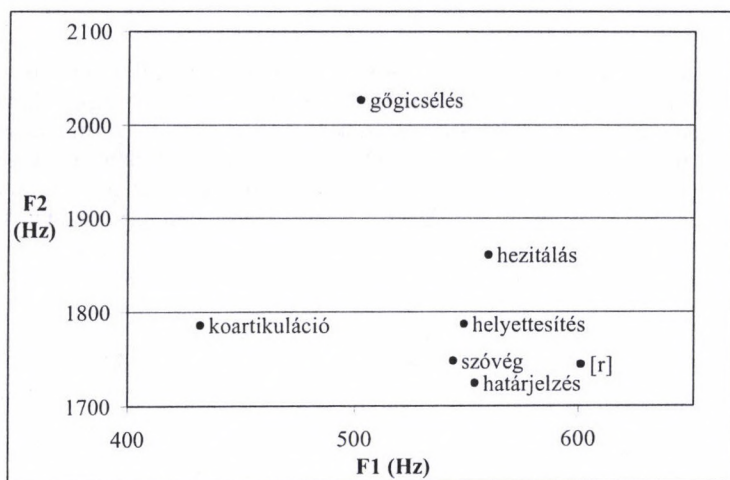
Kilencven olyan gögicsélt magánhangzó formánsszerkezetét elemeztük, amelyeket vagy svának, vagy [ø] magánhangzónak azonosított a jelen szerző hallás

alapján (a rögzített anyag többszöri meghallgatása után). A 3. táblázat a három formáns átlagértékeit és a szórásokat összegzi.

3. táblázat: Gögicsélt magánhangzók formánsszerkezete

| Magánhangzók | F ₁ (Hz) | | F ₂ (Hz) | | F ₃ (Hz) | |
|--------------|---------------------|-----------------|---------------------|-----------------|---------------------|-----------------|
| | Átlag | Átlagos eltérés | Átlag | Átlagos eltérés | Átlag | Átlagos eltérés |
| [ø] | 545,5 | 59,9 | 1985 | 273,9 | 3207,1 | 292,1 |
| [ə] | 502 | 47,85 | 2027 | 221,6 | 3021,4 | 242,3 |

A gögicsélt svák és [ø] magánhangzók első formánsai szignifikáns különbséget mutatnak ($t(89) = -2,409, p < 0,026$), a másik két formáns esetében ilyen eltérés nem volt igazolható. Ez egyfelől azt mutatja, hogy a hallás alapján elkülönített magánhangzók kétféle minőséget képviselnek; másfelől pedig hogy a csesemő a semleges magánhangzó mellett már képes a magyarra jellemző középső nyelvvállású, de a svától különböző képzésű magánhangzó ejtésére. A különbség artikulációsan a nyelv függőleges mozgásában jelentkezik. Ez utalhat arra, hogy a nyelv függőleges mozgása a szájüregben az anyanyelv-elsajátítás kezdetén elsődlegesebb a vízszintes mozgásnál.



11. ábra

A magyar semleges magánhangzók elhelyezkedése az első és a második formánsok függvényében

Láttuk, hogy az egyes nyelvekben a semleges magánhangzók képzése és formánsai – a szakirodalmi adatok alapján – óriási különbségeket mutatnak (2. áb-

ra). A magyar svá magánhangzók a különböző funkciókban szintén nagy akusztikai eltérésekkel realizálódnak, amiből arra következtetünk, hogy a képzésük is jelentősen eltér egymástól. Ha az összes mért semleges magánhangzó formánsainak átlagát nézzük, azt mondhatjuk, hogy a magyar svá inkább a szájüreg elülső részében, mint középen képzett. (Egyébként is tapasztalható a magyarban egyfajta tendencia a szájüreg elülső részének fokozottabb kihasználtságára.) Az [ø] magánhangzó formánsaitól a magyar semleges magánhangzó azonban egyértelműen elkülönül, ami azért hangsúlyozandó, mert ez a képzést tekintve támasztja alá a kettőjük közötti funkcionális különbséget (a svá nem fonémaértékű). Amíg az [ø] 380–460 Hz-es első, 1550–1650 Hz-es második és 2300–2450 Hz-es harmadik formánsokkal jellemezhető leggyakrabban, addig a svá jelen kíséreltben mért átlagértékei az F_1 -re 499 Hz, az F_2 -re 1760 Hz, az F_3 -ra pedig 2928 Hz-et mutatnak. Ha a magyar magánhangzórendszerben akarjuk a svát elhelyezni, akkor az valóban középponti helyet foglal el a szájüregben a képzése szerint, s ezt az objektív akusztikai adatok alátámasztják. A 11. ábra az F_1/F_2 koordináta-rendszerben ábrázolja a különböző funkciójú semleges magyar magánhangzók elhelyezkedését.

Következtetések

Az [ə] magánhangzót életünk első heteitől használjuk a legkülönbözőbb helyzetekben és funkciókban. Ez a centrális, multifunkcionális magánhangzó nagyon sok nyelvben fellelhető. A svá artikulációs konfigurációja elméletileg semleges, azonban az aktuális frekvenciaszerkezete nagy eltéréseket mutat az egyes nyelvek között, sőt egy adott nyelven belül is. A svá semlegessége – úgy tűnik – inkább percepciósi tény, nem annyira stabil magánhangzó, mint az összes többi. Az albániai toszk nyelvben például azt vették észre, hogy a svát egyre inkább a szájüreg hátsó részében képezik a beszélők (Granser–Moosmüller 2003). A holland mérések viszont azt erősítették meg, hogy a nyelvükben realizált svá relatíve stabil helyzetű a többi holland magánhangzó viszonylatában (van Bergem 1995). A magyar adatok azt mutatták, hogy a svá létezik a semlegesnek tekinthető ejtésben, ugyanakkor enyhe tendencia látszik az előrébb képzésre, vagyis a palatálisabb artikulációra. A svá semlegessége talán leginkább az észlelésben ragadható meg. A különböző nyelvek magánhangzóit az azon a nyelven nem tudó személy a saját percepciósi bázisán megsűrve azonosítja. Ebből arra következtethetünk, hogy más nyelvek semleges magánhangzóit hallgatva, például egy magyar anyanyelvű nem feltétlenül svát fog észlelni az elhangzásukkor. A „semlegesség” fonetikailag tehát csak azt jelenti, hogy mindig az adott nyelv magánhangzórendszerében értelmezve tekinthető a svá magánhangzó centrálisnak, semlegesnek vagy köz(ép)pontnak.

Több mint 200 évvel ezelőtt Kempelen azt írta a könyvében (1791), hogy az emberi fül könnyen „becsapható”, hiszen azt hallja, amit akar hallani (ti. észlelni). A semleges magánhangzó példája nyilvánvalóan igazolja ezt a megállapítást.

Irodalom

- Barry, William J. 1995. Schwa vs. schwa+/r/ in German. *Phonetica* 52. 228–235.
- Bergem, Dick van 1995. Perceptual and acoustic aspects of lexical vowel reduction, a sound change in progress. *Speech Communication* 16. 329–358.
- ten Bosch, Louis F. M. 1991. *On the structure of vowel systems. Aspects of an extended vowel model using effort and contrast*. Doctoral dissertation. University of Amsterdam, Amsterdam.
- Gósy Mária 1997. A magánhangzók minőségváltozásai a spontán beszédben. *Magyar Nyelvőr* 121. 45–56.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Granser, Theodor – Moosmüller, Sylvia 2003. Phonemic schwa – a challenge for Natural Phonology? The case of Albanian. In Katarzyna Dziubalska-Kolaczyk – Jarosław Weckwerth (eds.): *Future challenges for natural linguistics*. Lincom Europa, München, 165–178.
- Handbook 1999 = *Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet*. Cambridge University Press, Cambridge.
- Harmegnies, Bernard – Poch-Olivé, Dolores 1992. A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech Communication* 11. 429–437.
- Hirschfeld, Ursula – Wallraff, Uta 2002. Untersuchungen zum Schwa im deutschen. In Braun, Angelika – Masthoff, Herbert R. (eds.): *Phonetics and its applications*. Franz Steiner Verlag, Stuttgart, 493–505.
- Jones, Daniel 1966. *The pronunciation of English*. Cambridge University Press, Cambridge.
- Kerekes Ferentz 1836. *Értekezés és kitérések*. Debretzenben.
- Koopmans-van Beinum, Florens J. 1994. What's in a schwa? Durational and spectral analysis of natural continuous speech and diphones in Dutch. *Phonetica* 51. 68–80.
- Ladefoged, Peter – Maddieson, Ian 1996. *The sounds of the world's languages*. Blackwell, Oxford–Cambridge.
- Laver, John 1994. *Principles of phonetics*. Cambridge University Press, Cambridge.
- MacMahon, Michael K. C. 2003. Schwa. *The Phonetician* 87. 8.
- Onishi, Masao (ed.) 1981. *A grand dictionary of phonetics*. The Phonetic Society of Japan, Tokio.
- Pickett, James M. 1980. *The sounds of speech communication*. University Park Press, Baltimore.
- Sproat, Richard (ed.) 1998. *Multilingual Text-to-Speech synthesis: The Bell Labs approach*. Kluwer, Dordrecht, 209–212.
- Stevens, Kenneth N. 1972. The quantal nature of speech: Evidence from articulatory-acoustic data. In David, Edward E. – Denes, Peter B. (eds.): *Human communication: a unified view*. McGraw Hill, New York, 51–56.
- Vértes O. András 1980. *A magyar leíró hangtan története az újgrammatikusokig*. Akadémiai Kiadó, Budapest.
- Warner, Natasha L. – Jongman, Allard – Cutler, Anne – Mücke, Doris 2002. The phonological status of Dutch epenthetic schwa. *Phonology* 18. 387–420.

A KOARTIKULÁCIÓ A MAGÁNHANGZÓK TISZTA FÁZISA ÉS AZ ÁTMENETEK ÖSSZEFÜGGÉSÉBEN

Nikléczy Péter

Bevezetés

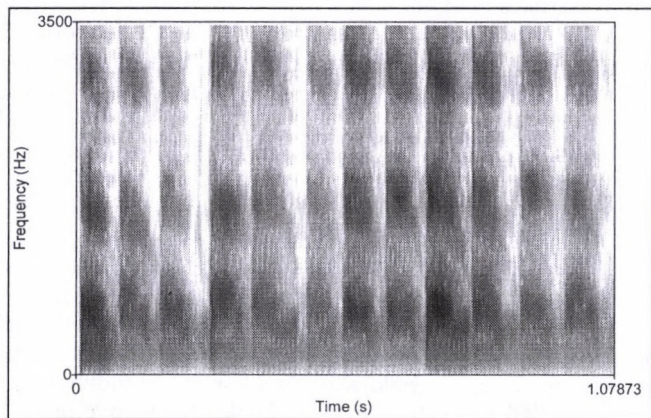
Folyamatos beszédkor a koartikuláció akusztikai következménye a magánhangzók formánsszerkezetének változása. Ez azt jelenti, hogy az adott hangra jellemző formánsstruktúra az adott hangkapcsolattól függően, különböző mértékben módosul. Rövid időtartamú beszédhang esetében a gyakorlatilag folyamatos artikulációs mozgás miatt szinte meghatározhatatlan az a következményes formánsszerkezet, amely voltaképpen reprezentálja az adott beszédhangot. A szakirodalmi adatok igazolják, hogy az aktuális kontextus nagymértékben befolyásolja a koartikuláció jelenségét, a fonetikai reprezentáció részben nyelvspecifikus, részben a hangképző szervek relatív tehetetlenségétől és nem kis mértékben egyéni ejtési sajátosságoktól is függ (Hardcastle–Hewlett 1999, Gósy 2004).

A jelen kutatás célja, hogy objektív választ kapjunk a koartikuláció mértékére a magánhangzó tiszta fázisának módosulásaiban. Azt elemeztük, hogy a magánhangzók tiszta fázisának frekvenciaingadozásai a hangkörnyezetből adódó következmények-e, avagy az egyén saját kiejtési módja befolyásolja a módosulásokat, változásokat. Az ejtésingadozások és a kontextuskülönbségek igen nagy variabilitása adatbázison történő vizsgálatokat fog igényelni, ehhez azonban nélkülözhetetlen az egyedi formánsmozgások sajátosságainak elemzése az időtényező tekintetbe vételével. A jelen megközelítés az alapoktól indul; akusztikai fonetikai elemzésekkel igyekeztünk a kérdésre választ kapni. A feltevésünk az, hogy a tiszta fázishoz képest az elemzett magánhangzók átmenetei jelentős eltéréseket mutatnak, amelyekben a kontextushatás meghatározóbb, mint az egyéni ejtésvariációk. Ha a hangkapcsolatban a magánhangzó tiszta fázisának formánsértékeit befolyásolja a hangkörnyezet, akkor minden azonos magánhangzó esetében bekövetkezik a tiszta fázis formánsstruktúrájának változása. Amennyiben ez nem mutatható ki, akkor a formánstruktúra változásáért nem a hangkörnyezet, hanem a beszélő változó artikulációja a felelős.

Kísérleti személyek, anyag, módszer

A kísérletekben négy magyar anyanyelvű nő vett részt, valamennyien pedagógusok, nemdohányzók (jelölésük a továbbiakban: „IL”, „IR”, „KR”, „M”). Életkoruk 25, 35, 45 és 55 év. A nyelvi anyag egyetlen mondat volt: *Sok szép já-*

téket mutattak be a szeptemberi előadáson. A félkövér betűk mutatják az elemzett magánhangzókat. A kísérleti személyek havonta egyszer elmondták a tesztmondatot, amelyet mindig ugyanazon körülmények között, laboratóriumban rögzítettünk. A mondatokat 10 kHz-es mintavételezési frekvenciával digitalizáltuk (az anyag és módszer megegyezett Gósy 2002 kísérletével). A mondatból két hangot vizsgáltunk meg ([ɔ, ɛ]) három különböző hangkapcsolatban. A hangokat a hangsorban elfoglalt helyükkel jelöltük (1a, 2a, 3a, 1e, 2e, 3e). Az eljárás a következő volt. Az azonos bemondótól származó, ugyanazon hangkapcsolatban lévő magánhangzókat (az átmeneti fázissal együtt) szegmentáltuk, és egymás után elhelyeztük egy adott hangfájlból (1. ábra). Így a tíz különféle időpontban rögzített azonos magánhangzó egymás mellé került. Ezt követően a magánhangzók tiszta fázisából kiemeltünk 4 periódust, majd a fentiek szerint ezeket is egymás mögé másoltuk Audacity 1.3 Beta hangeditáló programmal (2. ábra).

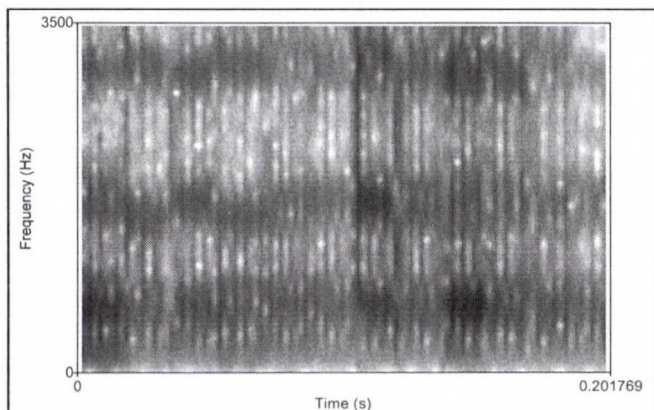


1. ábra

Az egymás mellé helyezett szegmentált magánhangzók az átmenetekkel

Az editált és összeszerkesztett hanganyagon spektrografikus elemzést végeztünk. Minden hang esetében megmértük a formánsok maximális és minimális értékét, valamint a hang tiszta fázisának formánsadatait. A számított átlagok alapján minden hangtípust kilenc érték reprezentált: F_1 -min, F_2 -min, F_3 -min; F_1 -max, F_2 -max, F_3 -max; F_1 -tf, F_2 -tf, F_3 -tf (ahol min = minimumérték, max = maximumérték, tf = tiszta fázis). Az értékeket minden esetben a hang kezdetén, a hang végén, illetve az audiovizuálisan megállapított tiszta fázis közepén mértük. Ezeket az értékeket magánhangzónként átlagoltuk, majd a maximum- és minimumpontok átlagának arányát adtuk meg százalékban. Tekintettel a tíz alkalomra, a két magánhangzóra és a 3-3 helyzetre, összesen 240 elemzést hajtottunk végre. Az akusztikai fonetikai vizsgálatokat a Kay CSL 4300B típusú digitális

jelanalizátorral, illetve a Praat 4.2 verziószámú programmal végeztük el. A statisztikai elemzéseket az SPSS 8.0 szoftverrel végeztük (95%-os szinten).



2. ábra

A szegmentált magánhangzók egymás mellé másolt tiszta fázisainak hangszíneképe

Eredmények

Az [ɔ]-ra vonatkozó adatokat az 1–4. táblázatokban összegeztük, a bemondókat a megfelelő betűjellel szemléltettük. A táblázatokban az 1a–3a jelölések a kontextusra utalnak; a legalsó sorban pedig a formánsok maximális és minimális értéke közötti különbség látható százalékban kifejezve.

1. táblázat: Az „IL” jelű beszélő [ɔ] magánhangzójának értékei (Hz)

| IL | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1a | 404 | 680 | 638 | 1533 | 1785 | 1670 | 2974 | 3166 | 3093 |
| 2a | 316 | 736 | 609 | 1379 | 1799 | 1670 | 2849 | 3133 | 2957 |
| 3a | 440 | 712 | 657 | 1428 | 1799 | 1582 | 2812 | 3059 | 2937 |
| Átl. | 386 | 709 | 634 | 1446 | 1794 | 1640 | 2878 | 3119 | 2995 |
| % | 52 | | | 20 | | | 18 | | |

2. táblázat: Az „IR” jelű beszélő [ɔ] magánhangzójának értékei (Hz)

| IR | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1a | 551 | 736 | 709 | 1405 | 1786 | 1543 | 2441 | 2960 | 2636 |
| 2a | 526 | 724 | 689 | 1465 | 1749 | 1564 | 2367 | 2873 | 2472 |
| 3a | 551 | 786 | 700 | 1305 | 1752 | 1399 | 2355 | 2700 | 2485 |
| Átl. | 542 | 748 | 699 | 1391 | 1762 | 1502 | 2387 | 2844 | 2531 |
| % | 38 | | | 22 | | | 17 | | |

3. táblázat: A „KR” jelű beszélő [ɔ] magánhangzójának értékei (Hz)

| KR | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1a | 588 | 872 | 733 | 1372 | 1762 | 1544 | 2478 | 2997 | 2663 |
| 2a | 489 | 749 | 658 | 1169 | 1712 | 1465 | 2256 | 2824 | 2559 |
| 3a | 526 | 971 | 743 | 1342 | 1725 | 1543 | 2515 | 3059 | 2661 |
| Átl. | 534 | 864 | 711 | 1294 | 1733 | 1517 | 2416 | 2960 | 2627 |
| % | 32 | | | 25 | | | 18 | | |

4. táblázat: Az „M” jelű beszélő [ɔ] magánhangzójának értékei (Hz)

| M | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1a | 502 | 736 | 681 | 1292 | 1589 | 1411 | 2676 | 2892 | 2780 |
| 2a | 539 | 699 | 681 | 1391 | 1725 | 1379 | 2490 | 2923 | 2710 |
| 3a | 539 | 736 | 695 | 1107 | 1589 | 1284 | 2602 | 2935 | 2719 |
| Átl. | 526 | 723 | 685 | 1263 | 1634 | 1358 | 2589 | 2916 | 2736 |
| % | 38 | | | 23 | | | 23 | | |

Az 5–8. táblázatok mutatják az „IL”, „IR”, „KR” és „M” jelű bemondók mondatbeli három [ɛ] hangjában mért akusztikai változásokat.

A mért adatok a vizsgált magánhangzók átlagos energiaszerkezetére vonatkozóan nem adnak felvilágosítást. Ennek az elemzésnek az elvégzésére a hosszú idejű analízis (LTA) nyújt lehetőséget. Az elkészített hangfájlok LTA-vizsgálata a formáns-energiamaximumokat veszi figyelembe, és a teljes „összefűzött” hangsor formánsszerkezetét az intenzitásviszonyok függvényében mutatja meg.

5. táblázat: Az „IL” jelű beszélő [ɛ] magánhangzójának értékei (Hz)

| IL | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1e | 476 | 740 | 642 | 1581 | 1917 | 1770 | 2758 | 3058 | 2897 |
| 2e | 488 | 668 | 631 | 1821 | 1989 | 1930 | 2962 | 3178 | 3071 |
| 3e | 500 | 632 | 620 | 1845 | 2135 | 2003 | 2914 | 3142 | 3026 |
| Átl. | 488 | 680 | 631 | 1749 | 2013 | 1901 | 2878 | 3126 | 2998 |
| % | 38 | | | 13 | | | 8 | | |

6. táblázat: Az „IR” jelű beszélő [ɛ] magánhangzójának értékei (Hz)

| IR | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1e | 464 | 680 | 604 | 1100 | 1905 | 1489 | 2457 | 2950 | 2601 |
| 2e | 512 | 752 | 683 | 1689 | 2049 | 1880 | 2746 | 2998 | 2867 |
| 3e | 548 | 728 | 700 | 1821 | 2097 | 1977 | 2649 | 3034 | 2874 |
| Átl. | 508 | 720 | 662 | 1536 | 2017 | 1782 | 2617 | 2994 | 2780 |
| % | 30 | | | 24 | | | 13 | | |

7. táblázat: Az „KR” jelű beszélő [ε] magánhangzójának értékei (Hz)

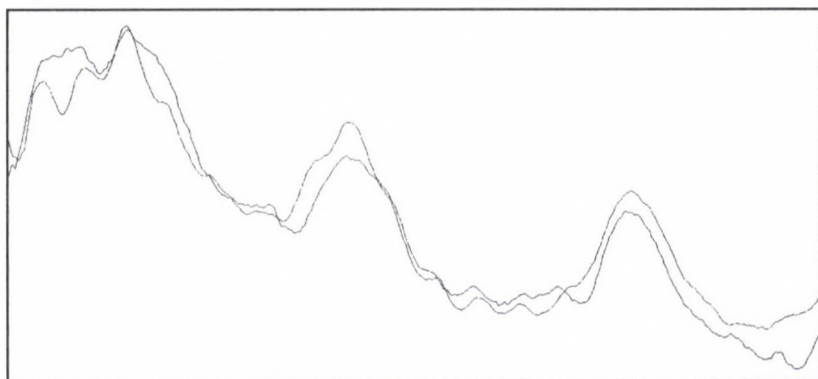
| KR | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1e | 524 | 836 | 729 | 1533 | 2145 | 1896 | 2469 | 3035 | 2761 |
| 2e | 512 | 716 | 693 | 1833 | 2157 | 2049 | 2649 | 3166 | 2903 |
| 3e | 548 | 740 | 693 | 1845 | 2193 | 2038 | 2770 | 3106 | 2918 |
| Átl. | 528 | 764 | 705 | 1743 | 2165 | 1994 | 2629 | 3102 | 2860 |
| % | 31 | | | 20 | | | 15 | | |

8. táblázat: Az „M” jelű beszélő [ε] magánhangzójának értékei (Hz)

| M | F ₁ -min | F ₁ -max | F ₁ -tf | F ₂ -min | F ₂ -max | F ₂ -tf | F ₃ -min | F ₃ -max | F ₃ -tf |
|------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|
| 1e | 441 | 766 | 592 | 1566 | 1977 | 1821 | 2566 | 2833 | 2718 |
| 2e | 428 | 680 | 512 | 1773 | 2049 | 1905 | 2697 | 2998 | 2819 |
| 3e | 452 | 644 | 621 | 1713 | 1989 | 1942 | 2661 | 2938 | 2892 |
| Átl. | 440 | 696 | 575 | 1704 | 2005 | 1889 | 2641 | 2923 | 2809 |
| % | 37 | | | 15 | | | 10 | | |

Ha az átmeneti fázist is tartalmazó összefüzdött hangsonon és a csak tiszta fázist tartalmazó hangsonon elvégezzük az LTA-elemzést, akkor a koartikulációs hatás esetében a két görbének különböznie kell. Ha a két görbe azonosnak tekinthető, akkor a koartikuláció nagy valószínűséggel nincs jelentős befolyással a tiszta fázis formánsszerkezetére (3. és 4. ábra).

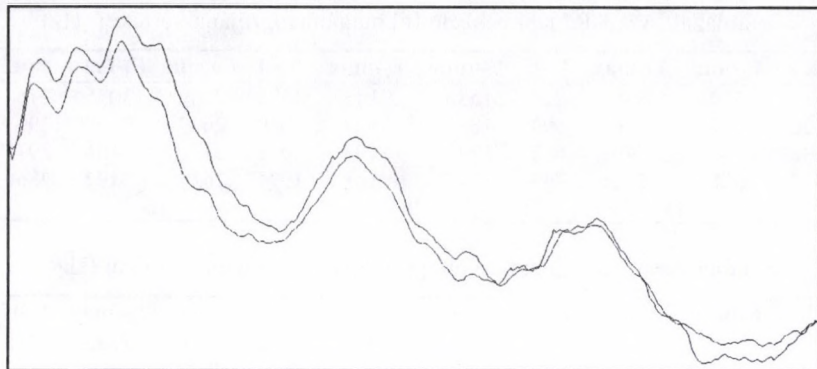
Az LTA-vizsgálatok eredményei numerikus formában a 9–16. táblázatból olvashatók le.



3. ábra

Az [ɛ] hang átlagolt energiaspektruma az „IL” jelű bemondó ejtése alapján.

A felső görbe az átmeneti fázist is tartalmazza, az alsó görbe csak a tiszta fázist



4. ábra

Az [ɛ] hang átlagolt energiaspektruma az „IL” jelű bemondó ejtése alapján. Az alsó görbe az átmeneti fázist is tartalmazza, a felső csak a tiszta fázist

9. táblázat: Az „IL” jelű beszélő [ɔ] magánhangzójának LTA-értékei (Hz)

| IL | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1a | 582 | 589 | 1687 | 1670 | 3056 | 3033 |
| 2a | 554 | 572 | 1650 | 1706 | 2967 | 2967 |
| 3a | 666 | 666 | 1594 | 1594 | 2891 | 2891 |
| Átl. | 600 | 609 | 1643 | 1656 | 2971 | 2963 |
| % | 2 | | 1 | | 1 | |

10. táblázat: Az „IR” jelű beszélő [ɔ] magánhangzójának LTA-értékei (Hz)

| IR | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1a | 530 | 547 | 1625 | 1502 | 2787 | 2658 |
| 2a | 634 | 634 | 1582 | 1602 | 2502 | 2562 |
| 3a | 662 | 710 | 1401 | 1421 | 2562 | 2514 |
| Átl. | 608 | 630 | 1536 | 1508 | 2617 | 2578 |
| % | 4 | | 2 | | 2 | |

11. táblázat: A „KR” jelű beszélő [ɔ] magánhangzójának LTA-értékei (Hz)

| KR | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1a | 763 | 779 | 1590 | 1606 | 2662 | 2662 |
| 2a | 671 | 649 | 1490 | 1458 | 2610 | 2590 |
| 3a | 691 | 680 | 1530 | 1575 | 2701 | 2638 |
| Átl. | 708 | 702 | 1536 | 1546 | 2657 | 2630 |
| % | 1 | | 1 | | 1 | |

12. táblázat: Az „M” jelű beszélő [ɔ] magánhangzójának LTA-értékei (Hz)

| M | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1a | 601 | 582 | 1437 | 1473 | 2763 | 2759 |
| 2a | 581 | 618 | 1494 | 1535 | 2722 | 2734 |
| 3a | 618 | 659 | 1182 | 1192 | 2702 | 2653 |
| Átl. | 600 | 619 | 1371 | 1400 | 2729 | 2715 |
| % | 3 | | 2 | | 1 | |

13. táblázat: Az „IL” jelű beszélő [ε] magánhangzójának LTA-értékei (Hz)

| IL | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1e | 585 | 585 | 1715 | 1715 | 2837 | 2830 |
| 2e | 578 | 578 | 1891 | 1875 | 3016 | 3023 |
| 3e | 538 | 538 | 1971 | 2020 | 2995 | 2987 |
| Átl. | 567 | 567 | 1859 | 1870 | 2950 | 2946 |
| % | 0 | | 1 | | 1 | |

14. táblázat: Az „IR” jelű beszélő [ε] magánhangzójának LTA-értékei (Hz)

| IR | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1e | 602 | 606 | 1718 | 1715 | 2694 | 2650 |
| 2e | 598 | 618 | 1867 | 1923 | 2803 | 2839 |
| 3e | 586 | 662 | 1971 | 1947 | 2795 | 2843 |
| Átl. | 595 | 628 | 1862 | 1861 | 2764 | 2777 |
| % | 5 | | 0 | | 1 | |

15. táblázat: A „KR” jelű beszélő [ε] magánhangzójának LTA-értékei (Hz)

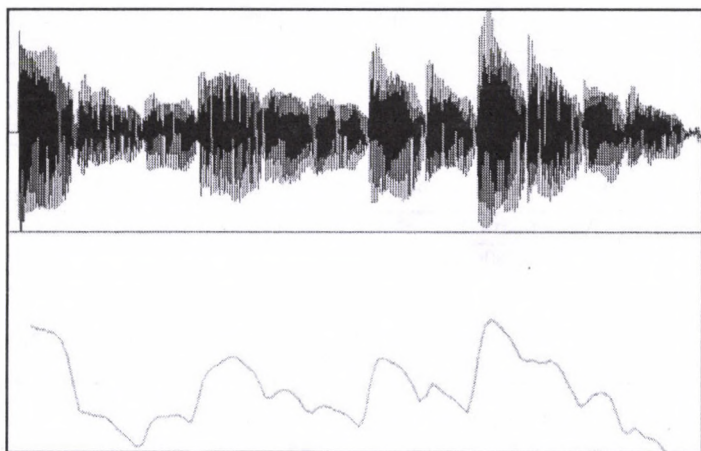
| IR | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1e | 658 | 742 | 1855 | 2000 | 2738 | 2775 |
| 2e | 642 | 642 | 2020 | 2020 | 2730 | 2935 |
| 3e | 630 | 591 | 1981 | 2059 | 2824 | 2857 |
| Átl. | 643 | 658 | 1952 | 2026 | 2764 | 2855 |
| % | 2 | | 4 | | 3 | |

16. táblázat: Az „M” jelű beszélő [ε] magánhangzójának LTA-értékei (Hz)

| IR | F ₁ LTA | F ₁ -tf LTA | F ₂ LTA | F ₂ -tf LTA | F ₃ LTA | F ₃ -tf LTA |
|------|--------------------|------------------------|--------------------|------------------------|--------------------|------------------------|
| 1e | 542 | 574 | 1815 | 1815 | 2694 | 2694 |
| 2e | 514 | 614 | 1891 | 1891 | 2799 | 2768 |
| 3e | 538 | 618 | 1891 | 1915 | 2867 | 2867 |
| Átl. | 531 | 602 | 1865 | 1873 | 2786 | 2776 |
| % | 12 | | 1 | | 1 | |

A táblázatokból kiolvasható, hogy az LTA-vizsgálatok eredménye jelentősen eltér a spektrografikus elemzésekkel kapott mérési adatoktól. A formánsok ma-

ximum- és minimumpontjainak mért értékei 8% és 52% között váltakoztak (1–8. táblázat), míg az LTA esetében ugyanezek az értékek 0% és 12% között ingadoztak. Ez azt jelenti tehát, hogy a koartikuláció hatása az elemzett magánhangzók és kontextusok esetében elhanyagolható. A kézenfekvő magyarázat az, hogy az elemzett magánhangzók belső intenzitás szerkezete ellensúlyozza a frekvenciamódosulások érvényesülését (5. ábra). A két vizsgált magánhangzóra kapott elsődleges mérési adatok alapján ugyanis egyértelműen kimutatható volt, hogy a hangok a tiszta fázisnál, avagy annak közelében érik el maximális intenzitásukat. Ennek következtében az átlagos energiaspektrumra a tiszta fázis kell, hogy döntő hatást gyakoroljon.



5. ábra

Az átmeneti fázist tartalmazó [ɔ] hangsor intenzitásgörbéje az „IL” jelű bemondó ejtésében

Választ vártunk arra a kérdésre is, hogy vajon az egyén artikulációja milyen hatással van a magánhangzó tiszta fázisára. A 17. táblázat a négy beszélő ejtésében mért tiszta fázisok formánsstruktúrát érintő százalékos eltérését mutatja. Megállapítható, hogy az [ɔ] esetében a legkisebb eltérés a beszélő személyek között az első formáns esetében van (11%), míg a legnagyobb a második formáns esetében (17%). Az F_3 értéke a kettő közötti ugyan, bár alig különbözik az F_2 -étől (16%). Az [ɛ]-nél viszont az F_1 esetében mértük a legnagyobb eltérést (19%), a legkisebbet pedig az F_3 -nál (8%). Az F_2 értéke ismét az F_3 -hoz adódott közelebb (11%). Összehasonlítva ezeket az értékeket az átlagolt energiaspektrumnál mért értékekkel, megállapítható, hogy a vizsgált két magánhangzó formánsértékeinek alakulására nagyobb hatással van az egyén ejtési módja, mint a hangra gyakorolt koartikulációs hatás.

17. táblázat: Az egyes bemondók közötti különbség a minimális és maximális F_1 , F_2 és F_3 számértéke alapján

| Magánhangzó | Maximális eltérés (%) | | |
|-------------|-----------------------|-------|-------|
| | F_1 | F_2 | F_3 |
| [ɔ] | 11 | 17 | 16 |
| [ɛ] | 19 | 11 | 8 |

Következtetések

A kutatás során elemzett két hang esetében kimutatható volt a hangok tiszta fázisára az egyéni artikuláció meghatározó szerepe. Az eredmény jelentősége elsősorban a beszélő személy kilétének meghatározásában van, ahol a rendelkezésre álló hanganyagok nem mindig tartalmaznak azonos hangkapcsolatokat. Természetesen szükséges további magánhangzók tiszta fázisának és hangátmeneteinek vizsgálata különféle kontextusokban, hogy a most megfogalmazott állítás valamennyi magánhangzóra igazolhatóan kiterjeszthető legyen.

Irodalom

- Gósy Mária 2002. Magánhangzók változása az idő függvényében. In Hunyady László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia 2002*. Debreceni Egyetem Kosuth Egyetemi Kiadója, Debrecen, 7–20.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Hardcastle, Wiliam J. – Hewlett, Nigel 1999. *Coarticulation*. Cambridge University Press, Cambridge.

A NAZÁLISOK OKOZTA SZERKEZETVÁLTOZÁS A ZÁR-, RÉS- ÉS ZÁR-RÉS HANGOKBAN MÁSSALHANGZÓ-KAPCSOLAT ESETÉN

Olaszy Gábor

Bevezetés

A beszédben a magánhangzók (V) és mássalhangzók (C) egyfajta váltakozása teszi lehetővé, hogy a percepció számára változatos, egymástól megkülönböztethető hangsorokat ejtsünk. A hangsorokban alapvetően négyféle hangkapcsolódási forma jöhet létre: VV, CV, VC és CC. Ezeknek sokféle további kombinációját használjuk a beszédben (Szende 1976). A beszédhangok hangsorba szerveződése során a hangkapcsolódás lényeges eleme a koartikuláció következtében létrejött úgynevezett átmeneti fázis, amely a beszédhang egy részét teszi ki. Az átmeneti fázis időben lezajló folyamat, a benne megvalósuló akusztikai tartalom folyamatosan változik, mivel az artikulációs szervek is folyamatosan változtatják helyzetüket (Hardcastle–Hewlett 1999). A változás az egyik hangra jellemző akusztikai tartalomról indul, és a hozzá kapcsolódó hangra jellemzőbe érkezik. Az átmeneti fázis akusztikai szerkezete a kapcsolatra jellemző, nem a kapcsolatban részt vevő hangokra. Akusztikai szinten a frekvenciaszerkezet változhat, a hang(ok) belső időszerkezete módosulhat és végül a hangkapcsolatra jellemző intenzitásmenet. A koartikuláció hatása tehát az, hogy e három elem valamelyike (vagy kombinációjuk) a legtöbb esetben változik a hangkapcsolatban a hangra egyébként jellemző akusztikai szerkezethez képest. A koartikulációra fordított energia függ a két hangra jellemző artikulációs képzési helytől, módtól és a gerjesztéstől, és ez áttételesen kifejeződik az akusztikai tartalomban is. Ha az artikulációs szerveket nem akadályozza semmi a koartikulációban, akkor azt mondjuk, hogy az átmenet simán jön létre. Ellenkező esetben mind az időtényezőben, mind az akusztikai tartalomban változások léphetnek fel. A magánhangzók közötti átmeneteknél például a nyelv, az ajkak, az állkapocs mozgása folyamatos. Mindhárom is változhat akadálytalanul, egymástól függetlenül, ezért a VV átmenet képzése koartikulációs szempontból simának mondható. A koartikulációt az sem bonyolítja, ha a nazális csatorna kinyílik. Ekkor a magánhangzó nazalizálódik (Földi 1989, Horváth 2005). A CV, VC kapcsolatoknál (Olaszy 1985) az átmenetre jellemző mozgássor bonyolultabb, mivel a mássalhangzókra jellemző képzési helyet és módot, valamint a két kapcsolódó hang közötti esetleges gerjesztésváltást is meg kell valósítani, de itt sem zavarják egymás mozgását különösebben az artikulációs szervek, a hangokra egyébként jellemző belső szerkezeti elemek nem torzulnak a kapcsolódáskor. Ha a nazális más-

salhangzó kapcsolódik a magánhangzóhoz, akkor az üregváltás simán végbe megy, a nazális üreget lezárjuk, és a hangképzés az orális csatornán keresztül folytatódik. Belátható, hogy a legbonyolultabb koartikulációs helyzet a két- és a többemeű mássalhangzó-kapcsolódásoknál lép fel, amikor az egyik hang képzési helyéből, módjából és gerjesztéséből kell a másik (esetleg a harmadik) kapcsolódó C képzésére jellemző artikulációs helyzetbe vezérelni a beszédszerveket. Ez bizonyos esetekben a mássalhangzó-kapcsolatban részt vevő hangok szerkezeti módosulásával is jár. Ebben a dolgozatban azt vizsgáljuk, hogy a nazálisok milyen hatást gyakorolnak a zár-, a rés- és a zár-rés hangok akusztikai szerkezetére a koartikuláció során.

Anyag, módszer, kísérleti személyek

Az akusztikai vizsgálatokat felolvasott szavakon végeztük. Az anyaggyűjtésnél az volt a szempont, hogy minden szó egy-egy mássalhangzó-kapcsolódást tartalmazzon. A vizsgálatok mérési anyagát egy többfunkciós, általunk készített szólista-adatbázisból kiválogatott kisebb mintaszókorpusz képezte (összesen 95 lexikai egység). Kétfajta kapcsolatot vizsgáltunk: CC-t és CCC-t, mindkettőt kizárólag szó belseji helyzetben. A CC kapcsolatoknál minden nazális hang és minden zár-, rés-, valamint zár-rés hang találkozására kijelöltünk egy-egy mintaszót. A CCC elemekre néhány, célzottan kiválasztott, nazális mássalhangzót is tartalmazó mintaszót vizsgáltunk. A szóállomány 10 bemondó felolvasásában állt rendelkezésre digitalizált formában 22 kHz/16 bit mintavételezéssel. A bemondók nem szerinti megoszlása 50%-os, életkoruk 30-60 év, foglalkozásukhoz szorosan hozzátartozik a beszéd mindennapi használata (bemondó, színész, egyetemi oktató). A hangfelvételt számítógépen tároltuk és készítettük elő a mérésekhez. Elvégeztük a szavak fonetikai, hangszintű átírását is, ez alapján bejelöltük a beszéd hullámformájában a hanghatárokat (címkézés). Az így kialakított vizsgálati anyag tehát minden mintaszóra a következő adatokat tartalmazta: a szó szöveges (karakteres) alakja, a szót alkotó hangsor hangjainak szimbólumai, a szó hullámformája, a hanghatárok címkéi. A felolvasott szólista átlagos artikulációs sebessége 10,5 hang/s. A közölt időtartam adatok minden esetben a 10 bemondó egyéni értékeiből képzett átlagot jelentik.

Mérési módszerek

A mérés módszerének kialakításánál a cél az volt, hogy a lehető legteljesebb képet adjuk a vizsgált mássalhangzó-kapcsolódások szerkezeti sajátosságairól. A vizsgálati eredményeket a VCV kapcsolatra jellemző mássalhangzó szerkezettel hasonlítottuk össze, ha ehhez képest volt változás, azt emeltük ki.

A hanghatárok kijelölése

A két- és többemeű mássalhangzó-kapcsolatok vizsgálatában a hanghatárok megállapítása az egyik legkritikusabb feladat, hiszen erre épülnek a hangidőtartam-adatok, valamint a spektrális kapcsolódási jellemzők mérései is. A számítógépes technika sokat segít a hanghatárok egyre pontosabb megállapításában, de az alapvető nehézségek ugyanúgy fennállnak, mint régen (van Santen 1992).

A hanghatárokat manuálisan jelöltük ki, vizsgálva a hullámformát, az intenzitás-görbét és esetenként a spektrogramot.

A mérések eszközei

A mérésekhez kétfajta szoftvert használtunk. A hullámformán végzett mérésekhez a Profidev elnevezésű, magyar fejlesztésű célszoftver (Olaszy et al. 2000) állt rendelkezésre, a hangspektrogramokat a Praat 4.0 szoftverrel állítottuk elő.

Hangjelölések

A számítógépes feldolgozáshoz saját hangjelölési rendszert alakítottunk ki a számítógépes ábrázolás egyszerűsítésére. Ebben a rendszerben a hangok nagy többsége megfelel a saját betűképének. Egyes mássalhangzóknál azonban nem a betűképet használjuk, hanem a saját, egykarakteres jelölést. Ezek a következők: $gy = G$, $ty = T$, $ny = N$, $sz = s$, $s = S$, $zs = Z$, $cs = C$. Egyes ábrákon és táblázatokban a hangokat ezekkel a szimbólumokkal adjuk meg, egyébként a hangok fonetikai jeleit használjuk szögletes zárójelök között.

Eredmények

A nazálisok okozta szerkezetváltozásokat a zár-, rés- és zár-rés hangokban a változás típusa szerint csoportosítottuk. A következő jellemző eltéréseket találtuk a VCV helyzetű mássalhangzók szerkezetével való összehasonlítások során.

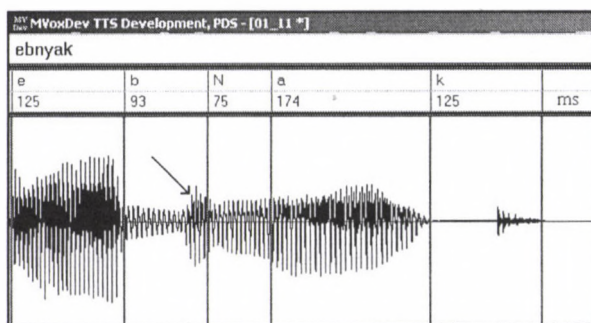
Svá töltelékelem képződése

A VCV kapcsolatok mássalhangzóinak szerkezetében nem találunk svá elemet. A svá mint töltelékelem megjelenése kifejezetten a mássalhangzó-kapcsolódásokra jellemző. A megjelenését az okozza, hogy az egyik képzési helyből a másikba való átmenetkor bonyolult artikulációs mozgásokat kell végrehajtani (Gósy 2006). A töltelékelem kialakulása a kapcsolódó mássalhangzók képzési helyének viszonyától függően változik.

Zárhangokban a svá töltelékelem akkor jöhet létre, amikor azok zöngés gerjesztésűek, és a CC kapcsolatban mint első elemek szerepelnek (*lábnak*, *ebnyak*). Ebben az esetben a zöngés zárhang zárfelpattanási része kapcsolódik a nazális hanghoz, és a svá a zöngés zárfelpattanást reprezentálja (1. ábra). Mely kapcsolatokra jellemző ez a jelenség? A választ az 1. táblázatban foglaltuk össze. Itt bemondóként jelöltük azokat a zöngés zárhangokat, amelyekben svá elemeket találtunk a nazális hanghoz való kapcsolódási pontokon.

A táblázatban a svák eloszlását önkéntesen két kategóriára bontottuk. Az egyik, amikor a bemondók többsége ejtette az elemet, ami azt jelentheti, hogy az ilyen kapcsolatokban a koartikuláció megvalósítására nagy energiát fordít a beszélő, artikulációs kényszerűségből jön létre a töltelékelem, tehát a két összekapcsolódó képzési helyre és módra jellemző a svá. A másik eset, amikor a bemondók kisebb része ejtette, ilyenkor inkább az artikulációs akarati határozza meg a svá létrejöttét, gondos ejtésnél, túlzott artikulációnál létrejöhet, lazánál nem jellemző. A magas F_2 -jú svávariáns létrejötte függ mind a zár-, mind a nazális hang képzési helyétől. Jellegzetesen a veláris zöngés zárhang és a palatális nazális zárhang találkozására jellemző ez a svávariáns. Ilyenkor az F_2 1700-

1800 Hz-re tolódik fel. Természetesen a keletkezett svák hossza és intenzitása is különbözik a fenti példákban, ezzel itt külön nem foglalkoztunk.



1. ábra

A [b] + [n] kapcsolat szerkezete az *ebnyak* mintaszóban

A függőleges vonalak a hanghatárok, a svát a nyíl mutatja.

A hangokat a számítógépes jelükkel jelöltük.

1. táblázat: A zöngés zárhangokban a zárpfattanást helyettesítő svá elem megvalósulása a nazálisokhoz való kapcsolódáskor
5 férfi (F1–5) és 5 női bemondó (N1–5) mintaszavaiban.

A * jelzés magas F₂-jú svávariánst jelöl.

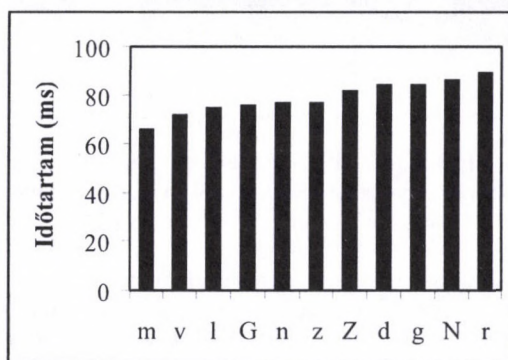
| Hangkapcsolat | Bemondó | | | | | | | | | |
|---------------|---------|-----|-----|-----|------|------|------|------|------|------|
| | N1 | N2 | N3 | N4 | N5 | F1 | F2 | F3 | F4 | F5 |
| [b] + [n] | svá | – | – | – | – | – | svá | – | svá | svá |
| [b] + [ɲ] | svá | svá | svá | svá | svá | svá | svá | – | svá | svá |
| [d] + [m] | svá | svá | svá | svá | svá | svá | svá | – | – | svá |
| [d] + [ɲ] | – | svá | svá | – | svá | – | svá | – | svá | – |
| [j] + [m] | svá | svá | – | – | svá | svá | svá | – | svá | svá |
| [g] + [n] | svá | svá | – | – | – | svá | svá | svá | svá | svá |
| [g] + [ɲ] | svá* | – | – | – | svá* | svá* | svá* | svá* | svá* | svá* |

Réshangok esetében a svát csak a [v] + [ɲ] kapcsolatokban találtunk, jellemzően a bemondók 50%-a ejtette. Ezek a svák intenzitásban gyengébbek, mint a zárhangoknál megvalósulóak. Ez is azt mutatja, hogy a [v] + [ɲ] kapcsolatnak nem lényegi eleme a réshang végén esetlegesen megjelenő svá.

A zárpfattanás elmaradása

A vizsgált C_{zár} + C_{naz} kapcsolatokban, amennyiben a két mássalhangzó képzési helye megegyezik, a zárhangok zárpfattanása nemigen mutatható ki. Ilyen esetek a [b p] + [m], a [d t] + [n] és a [j] + [ɲ]. A zárpfattanás elmaradása

kihat a mássalhangzó hosszának alakulására is, rövidülhet a hang, főleg a zöngés zárhangoknál. A zöngétleneknél ezt nemigen tudtuk kimutatni. A rövidülés ténye a más CC kapcsolatokkal való összehasonlításban látható szemléletesen. Példaként bemutatjuk az erre az összehasonlításra elvégzett mérésből származó [b] + C hangkapcsolatokból kapott hangidőtartam-átlagok diagramját (2. ábra). Az ábra adataiból látható, hogy a [b] időtartama az [m]-hez való kapcsolódásokban a legrövidebb, továbbá az is, hogy azokban a hangkapcsolatokban, amelyekben létrejön a svá elem a zárhang zárfelpattanásaként, a hang időtartama nyúlik (lásd: a [b] + [ɲ]-ben a [b] 87 ms).



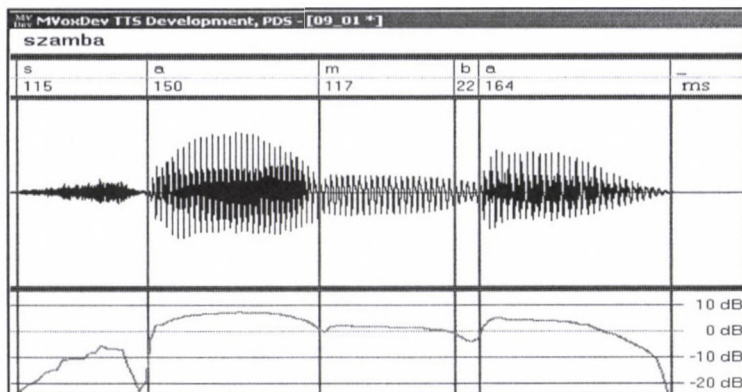
2. ábra

A [b] időtartama [b] + C hangkapcsolatokban a C függvényében
(A C-ket a számítógépes jelükkel jelöltük.)

Hangrészek rövidülése

A vizsgált hangkapcsolódásokban az artikuláció alakulásának következtében a zár- és zár-rés hangokban hangrészcsoportok rövidülések jöhetnek létre a nazális mássalhangzó hatására, illetve a hanghelyzet függvényében. Rövidülés akkor keletkezik, amikor a nazális mássalhangzó üregrendszerének nyitott állapota tulajdonképpen megzavarja az orális képzésű mássalhangzó kialakulását, azt a szerkezeti megvalósulást, amelyet VCV helyzetben megszoktunk (zárt orális üreg, tehát nincs hangkiáramlás és utána zárfelpattanás). A rövidülés (3. ábra) általában akkor jön létre, ha a nazális mássalhangzó a CC kapcsolat első eleme (*szamba, fénybe*).

A rövidülés annak köszönhető, hogy a zárhang képzésekor az orális zár ellenére a nazális üregen keresztül van folyamatos hangkiáramlás, így a tisztán orális ejtésre jellemző fojtott zöngé például a [b]-ben nem tud kialakulni, az üregváltás után szinte azonnal nyitjuk az orális zárat is. Ugyanakkor a [b]-t megelőző nazális mássalhangzó megnyúlhat, ahogy a 3. ábrán is láthatjuk.

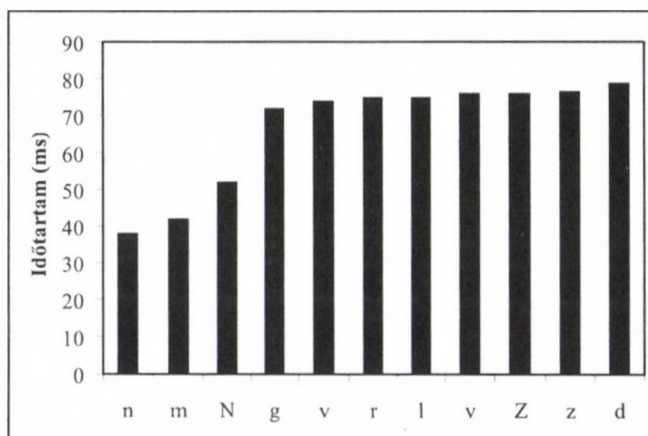


3. ábra

Példa a [b] zöngerészének rövidülésére
nazális mássalhangzó utáni helyzetben a *szamba* mintaszóban.

A teljes [b] hang csupán 22 ms-os

Ezek a rövidülések kihatnak a nazális hanghoz csatlakozó mássalhangzó teljes időtartamára is. Példaként bemutatjuk a C + [b] kapcsolatokra elvégzett mérésből származó hangidőtartam adatokat a 4. ábrán.

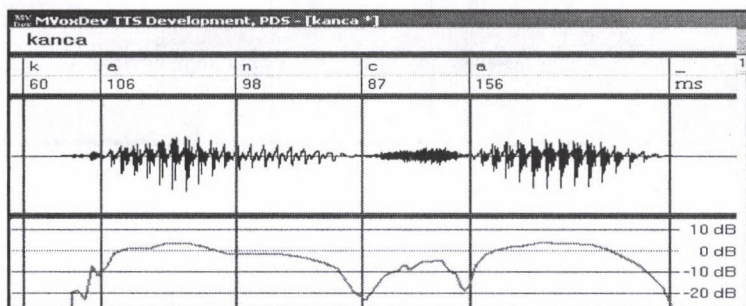


4. ábra

A [b] időtartama C + [b] hangkapcsolatokban a C függvényében
(A C-ket a számítógépes jelükkel jelöltük.)

A 4. ábra adataiból látható, hogy a [b] időtartama a nazálisok utáni helyzetben annyira lerövidülhet, hogy értéke mintegy fele az egyéb [b] + C kapcsola-

tokban mért értéknek. Találtunk olyan eseteket, zár-rés hangokban, amikor a rövidülés következtében a néma fázis teljesen elmarad. Erről Kovács (2002) is tesz említést. Ilyenkor a hangsor rezgéseképén a zár-rés hang úgy jelentkezik, mintha réshang lenne (5. ábra).



5. ábra

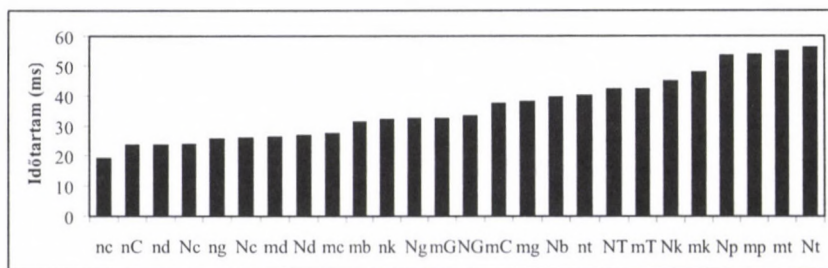
Példa arra, hogy a nazális hang hatására a zár-rés hangban eltűnhet a néma fázis a nagyfokú hangrészrövidülés következtében

Kérdés, hogy miért halljuk az ilyen hangrészt zár-rés hangnak. Erre egyrésztől az a magyarázat, hogy a szó lexikai tartalma behatárolja a lehetséges tartalmat (*kanca* szavunk nincs), másrésztől pedig az, hogy a réselem időtartama is meghatározó a percepcióban. A példában szereplő [ts] hang réseleme 87 ms-os, ugyanakkor az [s] hang időtartama 10,5 hang/s-os artikulációs sebességnél átlagosan 130 ms körüli (Olaszy 2006). A percepció rendszerrel szembeállítva egy automatikus hangazonosító szoftvert (például gépi beszédfelismerőben), a gépi rendszer viszont könnyen jelölheti réshangnak az ilyen hangrészeket. A rövidült zöngeszakaszok és néma fázisok időtartamait a 2. táblázatban összesítettük bemondónként. A táblázat adataiból származtattuk a mért zöngék és néma fázisok időtartamátlagait (6. ábra). Az ábrából kiolvasható tendenciák többek között a következők. A legrövidebb a zár-rés hangok néma fázisa az [n] után, hosszabb az [ŋ], majd a leghosszabb az [m] után.

2. táblázat: A zár- és zár-rés hangok zöngeszakaszainak és néma fázisainak időtartama (ms-ban) a nazális mássalhangzók utáni helyzetben öt női (N1–5) és öt férfi (F1–5) bemondó adataiból

| Hangkapcsolat | Bemondó | | | | | | | | | |
|---------------|---------|----|----|----|----|----|----|----|----|----|
| | N1 | N2 | N3 | N4 | N5 | F1 | F2 | F3 | F4 | F5 |
| [m] + [b] | 44 | 25 | 50 | 16 | 39 | 34 | 23 | 20 | 26 | 36 |
| [ŋ] + [b] | 25 | 32 | 60 | 22 | 45 | 35 | 32 | 29 | 54 | 61 |
| [m] + [p] | 55 | 53 | 56 | 52 | 52 | 48 | 54 | 50 | 68 | 49 |

| Hangkapcsolat | Bemondó | | | | | | | | | |
|---------------|---------|----|----|----|----|----|----|----|----|----|
| | N1 | N2 | N3 | N4 | N5 | F1 | F2 | F3 | F4 | F5 |
| [ŋ] + [p] | 57 | 61 | 67 | 55 | 52 | 35 | 59 | 40 | 69 | 40 |
| [m] + [d] | 36 | 26 | 18 | 28 | 29 | 22 | 21 | 26 | 25 | 32 |
| [n] + [d] | 31 | 23 | 26 | 17 | 13 | 23 | 22 | 21 | 29 | 31 |
| [ɲ] + [d] | 27 | 31 | 24 | 30 | 32 | 23 | 24 | 29 | 23 | 25 |
| [m] + [t] | 49 | 61 | 55 | 43 | 72 | 39 | 65 | 47 | 61 | 58 |
| [n] + [t] | 37 | 60 | 26 | 39 | 39 | 40 | 35 | 33 | 40 | 54 |
| [ɲ] + [t] | 46 | 50 | 49 | 63 | 71 | 44 | 43 | 56 | 74 | 66 |
| [m] + [g] | 32 | 38 | 31 | 31 | 47 | 33 | 35 | 45 | 59 | 31 |
| [n] + [g] | 23 | 33 | 36 | 22 | 15 | 22 | 23 | 19 | 23 | 41 |
| [ɲ] + [g] | 35 | 27 | 23 | 43 | 33 | 22 | 35 | 40 | 32 | 36 |
| [m] + [k] | 51 | 48 | 47 | 52 | 56 | 34 | 50 | 45 | 51 | 46 |
| [n] + [k] | 38 | 30 | 36 | 20 | 25 | 38 | 41 | 31 | 29 | 35 |
| [ɲ] + [k] | 36 | 35 | 45 | 57 | 55 | 40 | 49 | 36 | 55 | 41 |
| [m] + [j] | 35 | 34 | 31 | 23 | 34 | 34 | 33 | 32 | 29 | 41 |
| [ɲ] + [j] | 32 | 34 | 49 | 28 | 26 | 33 | 45 | 28 | 38 | 31 |
| [m] + [c] | 59 | 46 | 59 | 59 | 60 | 40 | 45 | 45 | 50 | 46 |
| [ɲ] + [c] | 39 | 46 | 34 | 52 | 55 | 45 | 24 | 33 | 45 | 51 |
| [m] + [t͡s] | 21 | 35 | 30 | 36 | 30 | 21 | 24 | 15 | 35 | 29 |
| [n] + [t͡s] | 10 | 30 | 20 | 17 | 25 | 0 | 29 | 20 | 20 | 20 |
| [ɲ] + [t͡s] | 30 | 26 | 24 | 58 | 31 | 0 | 15 | 20 | 31 | 25 |
| [m] + [t͡ʃ] | 30 | 51 | 30 | 41 | 37 | 30 | 51 | 25 | 40 | 40 |
| [n] + [t͡ʃ] | 29 | 20 | 29 | 30 | 25 | 15 | 11 | 25 | 21 | 30 |
| [ɲ] + [t͡ʃ] | 32 | 20 | 30 | 23 | 24 | 25 | 15 | 30 | 15 | 24 |



6. ábra

A nazálisok utáni zönge- és zárszakaszok időtartamai növekvő sorrendben (A vízszintes tengelyen a hangkapcsolat szerepel a hangok számítógépes jeleivel.)

A zöngés zárhangokban jobban érvényesül a rövidülés, mint a zöngétlenekben. Az [n] nagyobb rövidülést okoz, mint az [m p].

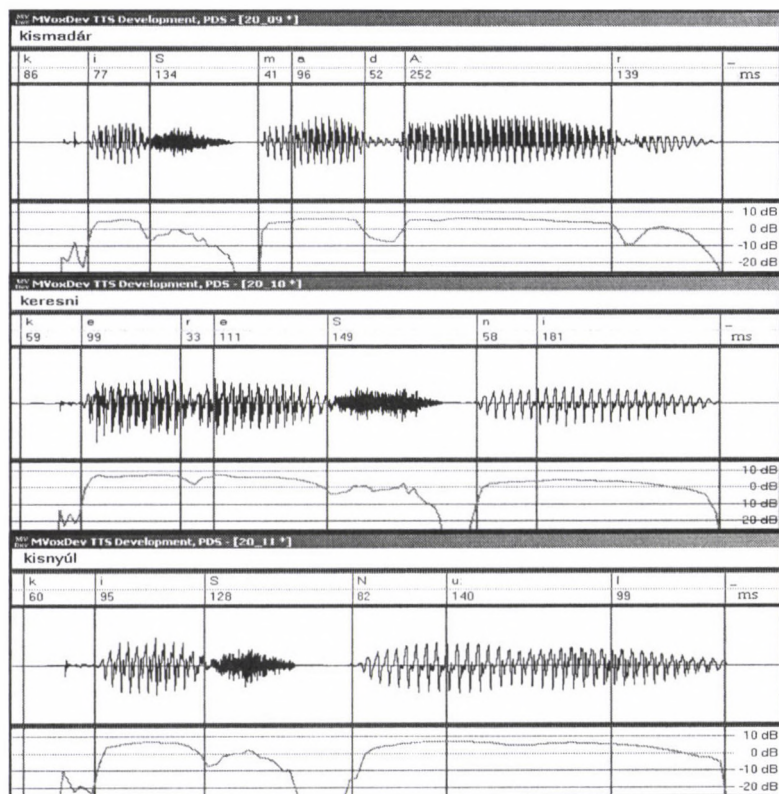
Másodlagos néma fázis kialakulása

A vizsgált kapcsolatokban kimutatható egy olyan 20-30 ms tartamú, kis intenzitású hangelem, amely a hangkapcsolódási ponton végbemenő gerjesztésváltáskor jön létre a nazális mássalhangzó zöngés rezgésének beindulása előtt. Ezt a szakaszt elneveztük **másodlagos néma fázis**nak. Vizsgálatáról irodalmi adatot nem találtunk.

Mi a magyarázata a jelenségnek? A nazális hangok képzésénél az orrüreg nyitott állapotban van, a levegő ezen keresztül áramlik, a hangszalagok rezegnek. Az orrüreg nyitását/zárását lényegileg a nyelvcsap végzi. Az üregváltáshoz (a nyelvcsap mozgatásához) idő kell. Ha a megelőző mássalhangzó zörejes gerjesztésű, akkor a gerjesztésváltáshoz is idő kell. A két időkomponens összeadódhat a zöngétlen mássalhangzók és a nazálisok találkozásánál, és ez okozza a másodlagos néma fázis kialakulását. Ennek a hangelemnek a tanulmányozása még további kutatásokat igényel. Itt csak bemutatjuk, hogy a hang hullámformarezgésképének vizsgálatánál milyen meglepő formációkkal találkozhat a kutató (vagy a gépi hangazonosító) az ilyen hangkapcsolatokban, továbbá kiindulási adatokat közlünk. Vegyük sorra zöngétlen hangonként a jelenséget. A zöngétlen zárhangok tekintetében a másodlagos néma fázis jelenléte nem is tűnik fel, egyrészt, mert maguk a zár-felpattanások is kis intenzitásúak, másrészt a zár-felpattanáshoz hozzá is tartozhat egy hosszabb lecsengő szakasz. Az elkülönítés tehát nehéz. Más a helyzet a zöngétlen réshangoknál és a zár-rés hangoknál. Itt a réselem intenzitása nagy, a rezgésképen tehát egyértelműen láthatóvá válik a másodlagos néma fázis a réselem után. A 7. ábrán látható másodlagos néma fázis jelenléte a hangsorban több kérdést is felvet. Először is kérdéses, hogy ez a hangszakasz melyik hanghoz tartozik, hogyan kell a hanghatárt ilyen esetekben bejelölni? Másodszor, hogy a rezgésképek alapján (főleg a *kisnyúl* mintaszóban) úgy látható, mintha a hangsorban egy zár-rés hang tükörképe szerepelne (tükör affrikáta?). Még szokatlanabb hangszerkezeti kép alakulhat ki ugyanezen hatás eredményeként a zár-rés hangoknál (8. ábra). Ekkor akár két néma fázis is lehet a zár-rés hangban, azzal kezdődik, és azzal is fejeződik be. A másodlagos néma fázisok időtartamait bemondónként a 3. táblázatban adjuk meg. A táblázat adatai azt mutatják, hogy az esetek nagy részében a bemondók megvalósították a másodlagos néma fázist (150 hangkapcsolatból 121-ben). Az is látható – ha a függőleges oszlopokat vizsgáljuk –, hogy bemondótól is függ a másodlagos néma fázis megvalósulása. Vannak olyanok, akik minden hangkapcsolatban képezték ezt a hangelemet, mások néha nem.

Hangonként is lehet összegezni az eredményeket. Leginkább az [s] + C_{naz} kapcsolatban jön létre a másodlagos néma fázis, utána következik az [ʃ], a [ts], a [tʃ], és végül legkevésbé a [f] + C_{naz} kapcsolatban. Ez a sorrend összefügghet a képzési helyekkel is. A másodlagos néma fázis jelenlétét megvizsgáltuk néhány híradásból származó felolvasott mondatban és spontán beszédben is.

Leginkább az [ʃ] + [m] hangkapcsolatra találtunk példákat a következő szakban: *elismételni, ismeretlen, ismeretlenek, ismerik, megismertük, megismer-tem, megismerkedtem*. Minden esetben kimutatható volt a másodlagos néma fázis a nazális hang rezgésének megindulása előtt.

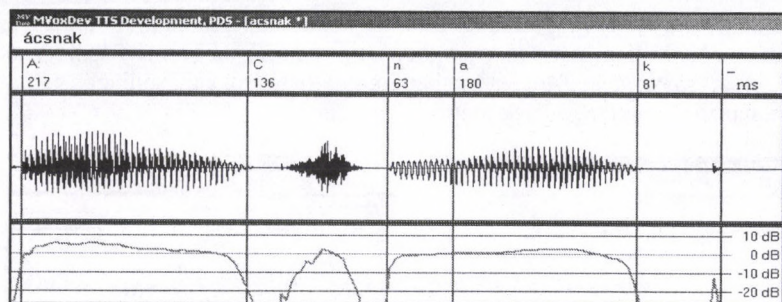


7. ábra

Példa az [ʃ] + C_{naz} hangkapcsolatban létrejövő másodlagos néma fázisokra a réshang és a nazális mássalhangzó találkozási pontjában
(A hanghatárjelöléseknél jelen vizsgálatban a néma fázist a réshanghoz számítottuk.)

Komplex hatások

A fenti hatások mindegyike érvényesül a CCC kapcsolatokban, sőt összegződhetnek is. A 8. ábrához hasonló rezgésképet kapunk például a *kulcsnál* szó zár-rés hangjára, a „tükör affrikáta” jelenik meg a *parancsnok*, a *harmincmillió*, a *harmincnyolc* szavak rezgésképeiben.



8. ábra

Az *ácsnak* szóban a fonológiai szintű zár-rés hang fizikai szerkezete a következő: zár, rés, zár, vagyis néma fázis, réselem, másodlagos néma fázis

3. táblázat: A másodlagos néma fázisok időtartamai ms-ban a zöngétlen rés- és zár-rés hangok végén a $C_{\text{rés}} + C_{\text{náz}}$ hangkapcsolatokban tíz bemondó produkciójában (N1–5 = női, F1–5 = férfi)

| Hangkapcsolat | Bemondó | | | | | | | | | |
|---------------|---------|----|----|----|----|----|----|----|----|----|
| | N1 | N2 | N3 | N4 | N5 | F1 | F2 | F3 | F4 | F5 |
| [f] + [m] | 23 | 20 | – | 30 | 25 | 20 | 15 | 20 | 34 | 20 |
| [f] + [n] | – | 15 | 20 | 20 | – | – | – | – | 15 | – |
| [f] + [ɲ] | 16 | 22 | 25 | 25 | – | 25 | 15 | – | 32 | 16 |
| [s] + [m] | 16 | 29 | 20 | 28 | – | 27 | 22 | 30 | 34 | 29 |
| [s] + [n] | 15 | 36 | 24 | 47 | 21 | 15 | 15 | 28 | 43 | 25 |
| [s] + [ɲ] | 21 | 32 | 35 | 31 | 20 | 28 | 20 | 29 | 25 | 25 |
| [ʃ] + [m] | 21 | 25 | 26 | 31 | 20 | 25 | 20 | 18 | 35 | 41 |
| [ʃ] + [n] | 19 | 30 | 30 | 40 | – | 20 | – | 24 | 25 | 21 |
| [ʃ] + [ɲ] | – | 36 | 36 | 34 | 24 | 27 | 31 | 20 | 45 | 25 |
| [t̪s] + [m] | – | 25 | 20 | 12 | 31 | 17 | – | – | 25 | 19 |
| [t̪s] + [n] | – | 25 | 25 | 25 | 25 | – | 20 | 25 | 25 | 27 |
| [t̪s] + [ɲ] | 25 | 25 | 20 | – | 25 | 33 | – | 20 | 25 | 25 |
| [t̪] + [m] | 17 | 19 | 20 | 30 | 15 | – | – | 25 | 20 | – |
| [t̪] + [n] | – | 17 | 25 | 40 | 25 | – | – | 15 | 25 | – |
| [t̪] + [ɲ] | 21 | 22 | 15 | 34 | 25 | – | 25 | – | 40 | 25 |

Összefoglalás

A cikkben bemutattuk, hogy a nazális mássalhangzók milyen szerkezeti változásokat okoznak a zár-, rés- és zár-rés hangokban mássalhangzó-kapcsolatokban. A vizsgálati eredmények szerint négy csoportba sorolhatók a szerkezeti vál-

tozások: svá elem megjelenése, zárpfelpattanás elmaradása, hangrészek rövidülése, másodlagos néma fázis kialakulása. Ez utóbbi hangelem szokatlan szerkezeti képet hoz létre (például a „tükör affrikáta” megjelenése a hangsorban). A másodlagos néma fázis nem illik bele a hangok szerkezeti leírásába, azonban fizikailag szerepel a hang képében. Jelenléte több kérdést is felvet, melyek megválaszolására további kísérletek és elméleti megfontolások szükségesek.

Irodalom

- Földi Éva 1989. Magánhangzó-nazalizáció, hosszú mássalhangzó vagy gemináta? *Egyetemi Fonetikai Füzetek* 2. 98–120.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2006. A semleges magánhangzó nyelvi funkciói. *Beszéd kutatás* 2006. 8–22.
- Kovács Magdolna 2002. Az affrikáták időszerkezetéről. In Hunyadi László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia 2002*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 39–54.
- Hardcastle, William J. – Hewlett, Nigel 1999. *Coarticulation*. Cambridge University Press, Cambridge.
- Horváth Viktória 2005. A magánhangzók nazalizációjáról. *Beszéd kutatás* 2005. 51–62.
- Olaszy Gábor 1985. *A magyar beszéd leggyakoribb hangsorépítő elemeinek szerkezete és szintézise*. Nyelvtudományi Értekezések 121. Akadémiai Kiadó, Budapest.
- Olaszy Gábor 2006. *Hangidőtartamok és időszerkezeti elemek a magyar beszédben*. Nyelvtudományi Értekezések 155. Akadémiai Kiadó, Budapest.
- Olaszy, Gábor – Kiss, Géza – Németh, Géza 2001. Hungarian audiovisual prosody composer and TTS development tool. In Puppel, Stanislaw – Grazina, Dermenko (eds.): *Prosody 2000*. Poznan, 167–178.
- Olaszy Gábor – Abari Kálmán 2005. Adatbázisok és számítógépprogramok a magyar beszéd időszerkezeti vizsgálatához. *Alkalmazott Nyelvtudomány V/1–2*. 41–62.
- Szende Tamás 1976. *A beszéd folyamat alaptényezői*. Akadémiai Kiadó, Budapest.
- van Santen, Jan P. H. 1992. Contextual effects on vowel duration. *Speech Communication* 11. 513–546.

Ezt a kutatást az OTKA T048429 támogatta.

KOARTIKULÁCIÓS FOLYAMATOK KÉT MAGÁNHANGZÓ KAPCSOLATÁBAN

Menyhárt Krisztina

Bevezetés

A magánhangzó-magánhangzó kapcsolatok vizsgálata hagyományosan a fonológia és a fonetika határterületére esik. Két magánhangzó találkozásakor fonológiai nézőpontból hiátus, míg fonetikai szemszögből (fonológiai) koartikuláció jön létre (vö. Gósy 2004, Siptár 2002a).

A fonológiában a hiátus két, szomszédos szótagmagot alkotó, magánhangzó egymásutánját, illetve a közöttük keletkezett ürt jelöli. Siptár (2002a) szerint a nyelvek ezt a helyzetet nem nagyon kedvelik, és igyekeznek legalább a fonetikai magvalósulás szintjén feloldani. Két magánhangzó találkozásakor a hiátushelyzetet a következőképpen lehet megszüntetni: *a)* az egyik vagy a másik magánhangzó törlődésével; *b)* egy mássalhangzó betoldásával; illetve *c)* az egyik magánhangzó szótagalkotó voltának megszűnésével (Siptár 2002b). Amennyiben egyik fent leírt eset sem következik be, akkor fonológiai szinten vagy fennmarad a hiátus, vagy egy olyan hang toldódik be, amelynek forrása a két szomszédos magánhangzó egyike. Ennek leggyakoribb formája a hiátustöltő [j] megjelenése a fonetikai kivitelezés szintjén. A [j] hang betoldása nem függ a morfémahatártól, a hiátus minden olyan esetben kitöltődik, amikor a hangkapcsolatban részt vevő magánhangzók egyike [i] vagy [i:]. Az [e:]-nél a kép változatosabb, a hangkapcsolat első tagjaként nem, míg második tagként opcionálisan okozhatja a hiátustöltő [j] megjelenését. A hiátushelyzet mindig fennmarad, amennyiben a hangkapcsolat hangjai alsó nyelvéllásúak és/vagy labiálisok (Siptár 2002a).

Fonetikai szempontból két magánhangzó találkozásakor (is) hangátmenet képződik, mint a koartikuláció akusztikai következménye, amely a beszéd folyamatosságát biztosítja. A hangátmenetnek több definíciója is lehetséges, figyelembe véve a hangkapcsolatot alkotó elemeket, illetve azt is, hogy akusztikai vagy percepciósi szemszögből vizsgáljuk-e az adott hangrealizációt. Így a hangátmenet *a)* olyan frekvenciamozgás, amely a magánhangzó kezdetétől a rá jellemző akusztikai pozíció bekövetkeztéig tart; *b)* az első hang tiszta fázisától a második hang tiszta fázisáig megfigyelhető formánsmozgás (ez vonatkoztatható két magánhangzó kapcsolatára); illetve *c)* a magánhangzó tiszta fázisától ugyanazon hang befejező pontjáig tartó formánsmozgás (Lehiste–Peterson 1992). Gósy (1995) percepciósi szempontú meghatározása szerint a hangátmenetek a koartikulációs jelenségek akusztikai következményei, amelyek a formánsmozgások révén a hangok minőségétől és kontextusától függően vagy invariáns je-

gyek, vagy másodlagos akusztikai kulcsok, vagy redundáns tényezőként funkcionálnak. Két magánhangzó találkozásakor a két hang közötti átmenet oly mértékben folyamatos lehet, hogy az artikuláció szempontjából akár diftongus is lehetne (Gósy 2004), a magyar köznyelvben azonban két egymást követő magánhangzó nem alkot diftongust, hanem külön szótagot képeznek. Bár néhány esetben folyik vita bizonyos magánhangzó-kapcsolatok (például az *autó* szóban) eltérő, azaz diftongusos voltáról (vö. Elekfi 1992; Kassai 1999), a köznyelvi ejtés akusztikai elemzése mégis azt mutatta, hogy más magánhangzó-kapcsolatokban is hasonló tendenciák figyelhetők meg, a különbség a frekvenciamódosulás mértékében és az időviszonyokban jelentkezik, amelyeket saját részről az egymás mellé kerülő, eltérő magánhangzó-minőségek határoznak meg (Gósy 2004). A magánhangzó-kapcsolódások időszerkezetét Olasz (1994) vizsgálta, eredményei szerint a magánhangzó-találkozások hangelemei hosszabb időtartamban valósulnak meg, mint ugyanazon hangok egyéb helyzetben. A hangátmenetek hossza akár a 80-100 ms-ot is elérheti, ami önmagában megfelel egy hangnak. A nyúlás hangsúlyosabban érvényesül a hosszú magánhangzóknál (vö. a *ráér* szóban). A szerző álláspontja szerint az elemek meghosszabbodása biztosítja a magánhangzó-kapcsolódás megfelelő hangzását.

A jelen dolgozat témája a magyar magánhangzó-kapcsolódások hangrealizációinak és koartikulációs folyamatainak vizsgálata, igényes köznyelvi ejtésben. A kutatás célja egyrészt új, objektív adatokkal jobban megismerni és feltárni a magánhangzó-kapcsolódások akusztikai szerkezetét (nagyobb korpuszon ilyen elemzés még nem történt), másrészt pedig, a kapott eredmények tükrében alátámasztani vagy korrigálni az eddigi fonológiai, illetve fonetikai kutatások megállapításait. Ehhez a következő kérdéseket foglalmaztuk meg: *a)* hogyan valósulnak meg a magánhangzók időtartam- és formánsszerkezet szempontjából egy magánhangzó-magánhangzó kapcsolat keretein belül; *b)* milyen koartikulációs jelenségek figyelhetők meg két magánhangzó kapcsolódásakor; és *c)* milyen a beszéd sebességének hatása a magánhangzó-kapcsolatok koartikulációs folyamataira.

Anyag és módszer, kísérleti személyek

A magánhangzó-kapcsolódások vizsgálatához összeállítottunk egy 40, egyszerű (2–4 szavas), értelmes magyar mondatból álló korpuszt (például: *Letört egy faág. Karambolozott egy autóbusz. A gyermek beilleszkedett a közösségbe.*). A vizsgálati anyag 40 hangkapcsolatot tartalmazott, amelyek viszonylag gyakran fordulnak elő a hétköznapi beszédben: *aa, au, aó, ai, ae, áí, áa, áe, áo, áé, áü, eo, ea, eu, eü, eő, ei, eé, éo, éá, iü, iő, iá, ió, iu, ia, ie, ié, oa, oá, öo, öe, óá, ői, uo, üé, üu, üe, üa.*

A kísérletben összesen 10 személy vett részt, öt nő és öt férfi. Mindnyájan felsőfokú végzettségűek voltak, és a köznyelvi standardot beszélték, beszédhibák nélkül. Életkoruk 23 és 60 év között mozgott, átlagosan 33 év volt. A kísérleti személyeket arra kértük, hogy szép kiejtéssel, átlagos tempóban egyszer olvas-

sák fel egymás után a negyven mondatot, a kísérlet céljáról előzetesen nem tájékoztattuk őket.

Az elhangzott közléseket hangszigetelt helységben, digitális jelfelvevőre rögzítettük, majd Praat 4.3 szoftverrel dolgoztuk fel. Meghatároztuk a beszélők artikulációs tempóját (hang/másodpercben), illetve folyamatos akusztikai és vizuális megerősítés mellett elemeztük az adott hangkapcsolat hangjainak időtartamát és formánsszerkezetét (első három formáns). Emellett adatoltuk a kapcsolat két tagja közötti esetleges koartikulációs jelenségeket. A kapott eredményeken leíró statisztikai elemzést, illetve korrelációanalízist végeztünk.

Eredmények

A kísérletben részt vevő személyeknél 11,86 hang/s-os átlagos artikulációs tempót mértünk, a leglassúbb beszélőnél 10,8, míg a leggyorsabbnál 12,9 hang/s-ot regisztráltunk (átlagos eltérés: 0,78), vagyis a mondatok felolvasása lassú/közepes artikulációs sebességgel történt.

A vizsgált magánhangzók időszerkezete és formánsértékei

Kísérleti anyagunkban összesen 12 különböző magánhangzó szerepelt a negyven hangkapcsolatban: [ɔ, a:, ɛ, e:, i, o, ɒ, ø:, u, y, y:]. Ezek időtartam-átlagait az 1. táblázat szemlélteti.

1. táblázat: A hangkapcsolatok hangjainak átlagos időtartama és szórásértékei az összes beszélőre

| Magánhangzó | Időtartam (ms) | | |
|-------------|----------------|--------------|-----------------|
| | Átlag | Határértékek | Átlagos eltérés |
| [ɔ] | 73,5 | 29–119 | 23,2 |
| [a:] | 109,2 | 52–197 | 27,8 |
| [ɛ] | 70,5 | 28–156 | 19,9 |
| [e:] | 85,8 | 52–137 | 17,3 |
| [i] | 69,1 | 36–114 | 17,1 |
| [o] | 82,1 | 35–140 | 20,3 |
| [o:] | 104,8 | 63–175 | 27,5 |
| [ø] | 71,2 | 38–104 | 19,1 |
| [ø:] | 90,7 | 41–159 | 24,8 |
| [u] | 68,5 | 30–168 | 25,9 |
| [y] | 72,4 | 31–125 | 22,1 |
| [y:] | 81,5 | 48–147 | 21,3 |

A kapott átlagadatok azt mutatják, hogy a nyelvileg rövid, illetve hosszú magánhangzók közül az utóbbiak hosszabb átlagos időtartamban valósulnak meg a magánhangzó-kapcsolatban. A legszűkebb határok között ez [e:] és az [i] időtartamának átlagos eltérése mozgott (17,1 és 17,3 ms), míg a legnagyobb szórásér-

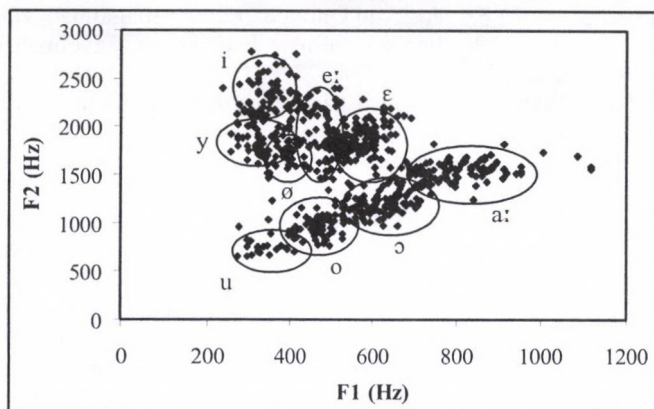
téket az [a:] és az [o:] magánhangzóknál figyeltük meg (27,8 és 27,5 ms). Mivel a jelen vizsgálathoz nem áll rendelkezésre olyan adathalmaz, amely az izolált szavakban előforduló magánhangzók időtartamára vonatkozna, ezért a mért értékeket a szakirodalomban közölt átlagos időtartamokkal vetettük össze (vö. Gósy 2004). A jelen kísérletben mért időtartamok alsó határértéke (az [e:] kivételével) minden esetben rövidebb, mint a szakirodalmi átlagoké. Ez több okra vezethető vissza, mint a beszédanyag különbözősége, a magánhangzók közötti koartikulációs folyamatok eltérései, de nem hagyhatjuk figyelmen kívül az egyéni ejtési sajátosságokat sem.

A magánhangzók formánsértékeit a 2. táblázat szemlélteti. A szakirodalmi adatokkal történt összehasonlítás a formánsértékek esetében is tájékoztató jellegű, az azonban megállapítható, hogy a saját mért adataink közelítik a szakirodalomban megadott formánsátlagokat (pl. Vértes 1982; Gósy 2004). A formánsértékek átlagos eltérése az F_1 -nél 32,6 és 98,4 Hz közötti, a szórás a magánhangzók többségénél a 40 Hz és 50 Hz közötti tartományba esett. Kivétel az [y], amelynél a legalacsonyabb, illetve az [e:] és az [a:], amelyeknél a legmagasabb átlagos eltérést mértük. Az F_2 -nél a magánhangzók szórására 103,6 Hz és 206,0 Hz közötti értékeket kaptunk. Legkisebb mértékben az [e:] és az [a:], legnagyobb mértékben pedig az [i, ø, y] értékei tértek el az átlagostól. A harmadik formáns átlagos eltérése 137,8 és 231,0 Hz között jelentkezett, a két szélső érték az [y], illetve az [i] hangnál volt mérhető. A többi magánhangzó szórása viszonylag szűk határok között mozgott, 140 Hz és 160 Hz között alakult.

2. táblázat: A hangkapcsolatok hangjainak átlagos formánsértékei az összes beszélőre

| Magánhangzók | Formánsok (Hz) | | | | | |
|--------------|----------------|-----------------|-------|-----------------|-------|-----------------|
| | Átlag | F_1 | Átlag | F_2 | Átlag | F_3 |
| | | Átlagos eltérés | | Átlagos eltérés | | Átlagos eltérés |
| [ɔ] | 613 | 55,1 | 1206 | 127,3 | 2455 | 165,1 |
| [a:] | 798 | 98,4 | 1526 | 104,9 | 2520 | 149,2 |
| [ɛ] | 559 | 53,1 | 1872 | 148,4 | 2629 | 184,4 |
| [e:] | 442 | 61,2 | 2203 | 103,6 | 2834 | 151,1 |
| [i] | 340 | 40,2 | 2313 | 200,4 | 2913 | 231,3 |
| [o] | 467 | 46,9 | 950 | 114,4 | 2446 | 169,9 |
| [o:] | 442 | 51,6 | 905 | 138,5 | 2405 | 151,4 |
| [ø] | 464 | 40,6 | 1594 | 206,6 | 2399 | 160,7 |
| [o:] | 414 | 45,5 | 1730 | 150,2 | 2499 | 153,4 |
| [u] | 341 | 42,1 | 769 | 113,4 | 2355 | 166,7 |
| [y] | 345 | 32,6 | 1713 | 186,6 | 2345 | 137,8 |
| [y:] | 328 | 49,3 | 1866 | 139,3 | 2335 | 147,5 |

Az első és a második formáns viszonyát az 1. ábrán szemléltettjük az összes beszélő vizsgált magánhangzóinál, ahol a sűrűsödési ellipszisek az F_1 és F_2 al-kotta koordináta-rendszerben az adott magánhangzó formánsainak frekvenciaértékeit tartalmazzák. A tíz beszélő adataiban, a nyilvánvaló átfedésekkel ugyan, a különböző magánhangzók jól elkülöníthetők egymástól.



1. ábra

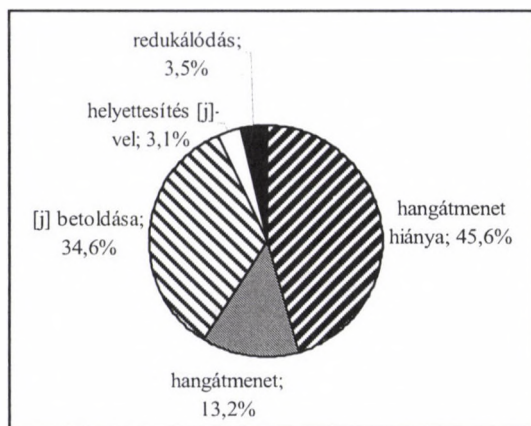
Az F_1/F_2 eloszlás a 10 beszélő összes vizsgált magánhangzójánál

12 esetben (az összes kapcsolódás 3%-a) előfordult, hogy a hangkapcsolat egyik tagja helyett egy másik, rendszerint hozzá képzésben közel álló hangra jellemző idő- és formánsstruktúrát regisztráltunk. Például az [εo]-ból [ɔo] és [øo], az [y:u]-ból [iu, eu] lett, illetve leggyakrabban az [y] és az [ø] váltakozását figyeltük meg: [a:y] → [a:ø]; [εy] → [εø]; [y:ɔ] → [øɔ]; [øa:] → [y:a:]; [ø:i] → [y:i]. Feltelezhető, hogy e jelenség okát az egyes beszélők sajátos ejtésében kereshetjük.

A magánhangzó-kapcsolatok koartikulációs folyamatai

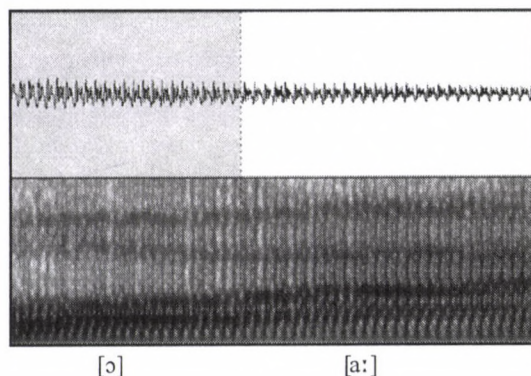
A jelen kísérletben a következő koartikulációs jelenségeket adatoltuk két magánhangzó kapcsolatában (2. ábra). Az esetek 45,6%-ában a két magánhangzó között nem regisztráltunk hangátmenetet, míg 13,2%-ban a két hang között megjelent a hangátmeneti szakasz. A [j] betoldása a két magánhangzó között 34,6%-ban volt megfigyelhető, míg az egyik magánhangzó [j]-vel történő helyettesítése 3,1%-ban jelent meg. 3,5%-ban fordult elő az, hogy a hangkapcsolat egyetlen hangban realizálódott.

A **hangátmenet hiányán** azokat az eseteket értjük, amikor az első hang utolsó, lecsengő periódusa után a következő hang első periódusa kezdődik, és a két magánhangzó folyamatos audiovizuális ellenőrzés mellett pontosan elkülöníthető egymástól (3. ábra). Mind az oscillogramon, mind a spektrogramon jól látható az [ɔ] hang vége és az őt követő [a:] hang kezdete.



2. ábra

A koartikulációs jelenségek százalékos megoszlása

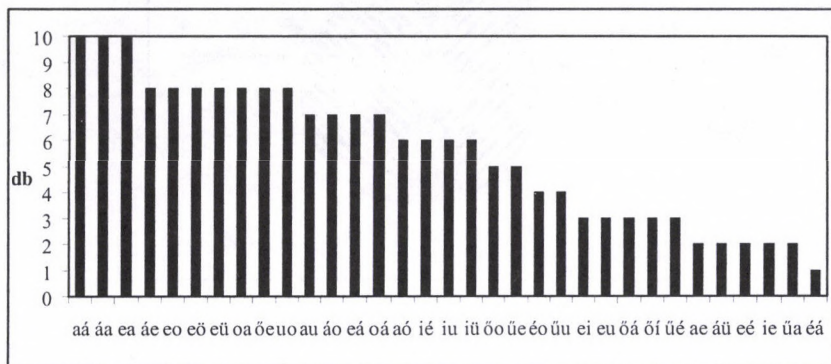


3. ábra

Hangátmenet hiánya az [ɔa:] hangkapcsolatban a *Leesett egy faág.* példamondatból, női ejtésben (a kijelölt terület az [ɔ] hang)

A hangátmenet hiányát az összes elemzett hangkapcsolat 85%-ánál regisztráltuk, de az előfordulási gyakoriság hangkapcsolatonként eltérő volt (4. ábra), például az *aá*, *áa*, illetve *ea* kapcsolódásoknál egyik beszélőnél sem volt hangátmenet, addig az *éa*-nál ez csak egyszer fordult elő. Ha összevetjük a kapott adatokat a fonológiai szempontú szakirodalmi megállapításokkal (Siptár 2002a), látható, hogy bár állításaik tendenciaszinten érvényesülnek, a fonetikai megvalósulásnál sokkal színesebb kép rajzolódik ki. Az alsó nyelvállású és/vagy labiális hangokból felépülő kapcsolódásoknál 73,2%-ban figyeltük meg a hangátmenet nélküli realizációt. Az [i] hangot tartalmazó kapcsolódások 14,1%-ánál nem volt hang-

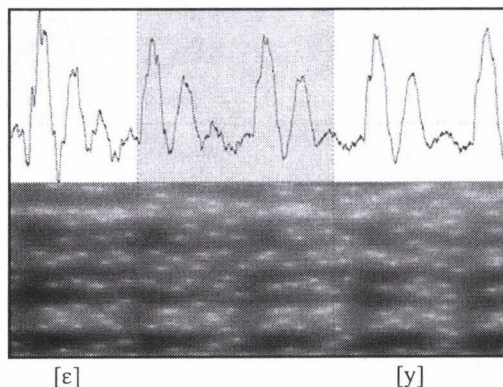
átmenet, például az *ié*, *iu* és *iü* a tíz beszélőből hatnál, az *ei* és az *ői* pedig háromnál hangátmenet nélkül valósult meg. Az [e:] -nél, amennyiben az a hangkapcsolat első pozíciójában volt, 25%-ban, második tagként 30,5%-ban szintén nem volt hangátmenet. Ezek a tények tovább árnyalják a korábbi elméleti leírásokat.



4. ábra

A hangátmenetek hiányának gyakorisági mutatói hangkapcsolatonként

Az artikulációs szervek relatív tehetetlensége folytán keletkeznek hangátmenetek. A jelen esetben ezek olyan formánsmozgások, amelyek a két magánhangzó tiszta fázisai között valósulnak meg, és audiovizuálisan jól elkülöníthetők (5. ábra). A magánhangzók percepcióját ezek az átmenetek nem befolyásolják, ezért Gósy (1995) meghatározása szerint az észlelés szemszögéből szükségtelenek.

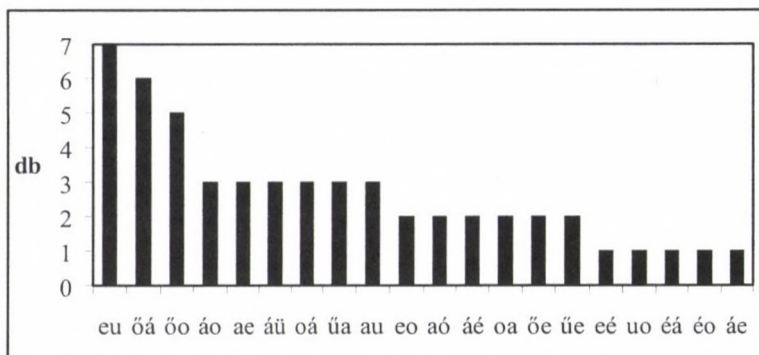


5. ábra

Hangátmenet megjelenése (a kijelölt terület) az [εy] hangkapcsolatban
A néni beült a buszba. példamondatban, férfi ejtésben

A jelen kísérletben a tíz beszélőtől származó összes adat 13,2%-ánál regisztráltunk hangátmenetet. Átlagos időtartamuk 14,9 ms volt (határértékek: 9–22 ms), vagyis ezek az átmenetek 1–3 periódust foglalnak magukba. Az adatok átlagos eltérése 3,04 ms, ami azt mutatja, hogy az átmenetek a hangkörnyezettől relatíve függetlenül, hasonló időtartamban valósulnak meg.

Hangátmeneteket az összesen 40 hangkapcsolat 50%-ánál figyeltünk meg (6. ábra). Ha a kapcsolódásokban részt vevő hangok szemszögéből nézzük a megjelenési arányokat, megállapíthatjuk, hogy hangátmenetek az alsó nyelvállású, illetve labiális hangokból álló kapcsolódások, továbbá a felsoroltak és az [ε] kombinációinál fordulnak elő. Az alsó nyelvállású és/vagy labiális magánhangzók közül felépülő hangkapcsolatok esetében a hangátmenetek gyakorisága 25,8%-os volt. Az [e:] -t tartalmazó kapcsolatokban 8,9%-ban fordult elő, míg az [i] kapcsolataiban egyáltalán nem találtunk hangátmenetet.



6. ábra

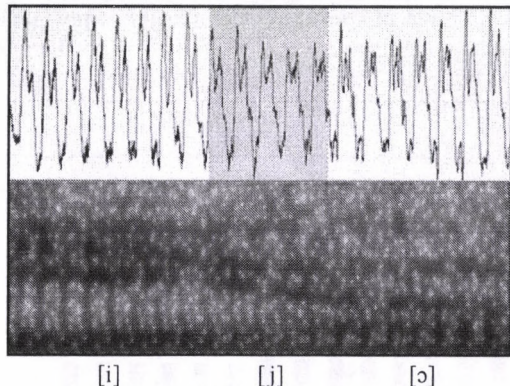
A hangátmenetek megjelenési arányai hangkapcsolatonként

Az adatok másik nagy csoportjánál a két magánhangzó között [j] **hang betoldását** figyeltük meg. A 7. ábrán látható a [j] megjelenése a két magánhangzó között, időtartama a példában 34 ms.

Megvizsgáltuk a két magánhangzó között ejtett [j] akusztikai szerkezetét és időtartamát (3. táblázat), majd összehasonlítottuk az önálló mássalhangzó pozíciójában lévő [j]-re vonatkozó szakirodalmi adatokkal. Megállapítottuk, hogy a betoldott [j] átlagosan mintegy 20 ms-mal rövidebb időtartamban valósul meg, mint a fonémaértékű [j] (vö. az Olaszky 2001 és a Menyhárt 2003 által megadott 59 ms-os, illetve 52,6 ms-os átlagértékekkel). Az időtartam átlagos eltérése azt mutatja, hogy a két magánhangzó közötti betoldásként megvalósuló [j] mássalhangzó relatíve hasonló időintervallumokban jelenik meg (a leggyakoribb érték a 25–45 ms).

A betoldott [j] formánsstruktúrája, összehasonlítva az önálló mássalhangzóként létrejövő [j] hangéval, szélesebb határok közötti megvalósulási értékeket

mutat. A korábbi kutatások adatai a [j] első formánsának határértékeit 300 Hz és 600 Hz közé teszik, a második formánsét pedig 1700 Hz és 2900 Hz közé, (az átlagértékek 413 Hz és 2227 Hz), míg a harmadik formáns átlagértéke 3458 Hz (Menyhárt 2003). A jelen kutatás eredményei a fent megadott adatoktól valamennyire eltérnek, ez a különbség az F_3 esetében a legszembetűnőbb. A különbség főként a betoldott [j] helyzetéből eredeztethető; feltételezzük, hogy az sokkal jobban függ a környező magánhangzók minőségétől, mivel a képzésük során működő koartikulációs folyamatok következtében jön létre.



7. ábra

[j] hang betoldása az [io] hangkapcsolatban, a *Kérlek, ne kiabálj.* példamondatból, női ejtésben

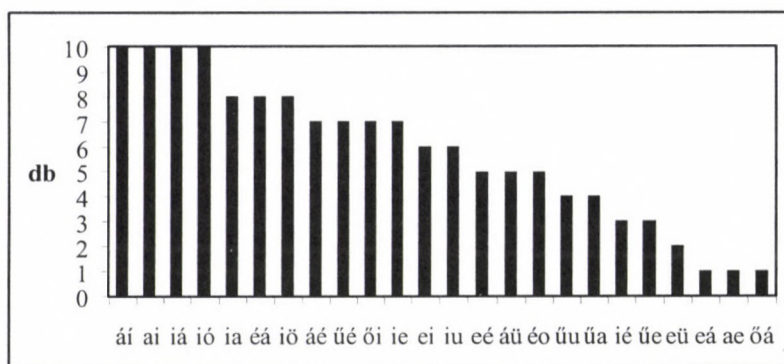
3. táblázat: A betoldott [j] időtartama és formánsszerkezete

| [j] hang | Időtartam (ms) | F_1 (Hz) | F_2 (Hz) | F_3 (Hz) |
|-----------------|----------------|------------|------------|------------|
| Átlag | 35,2 | 445 | 2071 | 2710 |
| Határértékek | 14,3–65,5 | 290–807 | 1242–2774 | 2127–3445 |
| Átlagos eltérés | 10,53 | 99,4 | 296,7 | 285,5 |

A [j] megjelenését az összes hangkapcsolat 60%-ánál regisztráltuk, hangkapcsolatonként eltérő arányban (8. ábra). Négy olyan kapcsolódás volt – az *ai*, *ai*, *ia*, *io* –, amelyeknél minden esetben megvalósult a [j] betoldása, legritkábban (csak egy-egy beszélőnél) pedig az *eá*, *ae* és *óa* hangkapcsolatoknál jelent meg. Ez utóbbiakat leszámítva megállapíthatjuk, hogy [j] megjelenését négy hang – [i, e:, y, y:] idézte elő. Legnagyobb arányban az [i]-nél fordult elő: 64,2%-ban, ezt követte az [e:] 19,6%-kal, és az [y, y:] összesen 13,6%-kal.

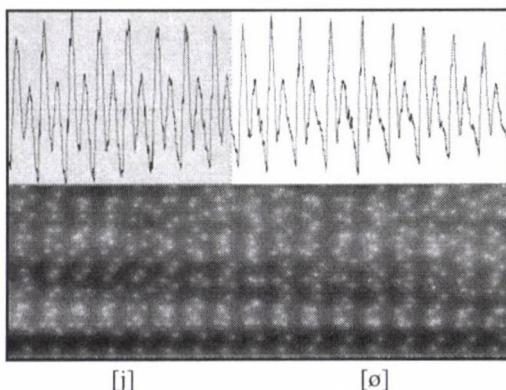
A fonológiai szakirodalom szerint a [j] betoldása a fonetikai megvalósulás szintjén [i] mellett mindig realizálódik, míg az [e:] hangnál függ annak helyétől a hangkapcsolaton belül, illetve a mellette lévő magánhangzó minőségétől (Sip-

tár 2002b). Ha a jelen vizsgálat összes [i]-t tartalmazó magánhangzó-kapcsolatát nézzük, a [j] betoldása 74,7%-ban valósult meg. Ez az arány az [e:]-nél is magas, 66,7%, emellett a hangkapcsolatban elfoglalt helyétől függetlenül jelent meg. A kapcsolódás első hangjaként az [e:] 65,0%-ban, míg másodikként 66,7%-ban idézte elő a [j] mássalhangzó realizálódását. Noha a labiális hangok környezetében betoldott [j] hang megjelenését nem vártuk, mégis 35,4%-ban az [y]-t, illetve [y:] -t magába foglaló kapcsolódásoknál is megjelent (itt nem szerepelnek az [e:] -t és [i] -t is tartalmazó adatok).



8. ábra

A betoldott [j] megjelenési gyakorisága hangkapcsolatonkénti bontásban

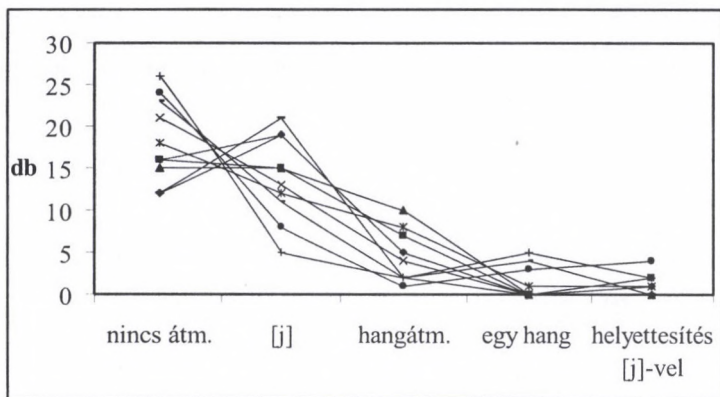


9. ábra

[j] hang (kijelölt rész) megjelenése [i] helyett az [iö] hangkapcsolatban a *Kiöblítettem a lavórt.* példamondatból, női ejtésben

A hangrealizáció szintjén 3,1%-ban jelent meg a hangkapcsolat egyik magánhangzójának **helyettesítése [j] mássalhangzóval** (9. ábra). Ez leginkább az [i] hangot érintette, és összesen 10 esetben fordult elő, az *ia, ie, iö, iu, iü* hangkapcsolatokban. Kétszer találtunk [j]-t az [y] hang helyén az *üu* kapcsolódásnál a *műugrók* [mjugro:k] szóban. A vizsgált pozícióban a [j] átlagos időtartama 47,96 ms (átlagos eltérés 18,9), vagyis hosszabb, mint a betoldott [j] hangé (vö. 3. táblázat). A formánsok átlagai a következőképpen alakultak: $F_1 = 368$ Hz, $F_2 = 2250$ Hz és $F_3 = 2839$ Hz. Feltételezhetjük, hogy az eredeti, hosszabb időtartamban realizálódó magánhangzónak a rövidebb időtartamú [j]-vel történő helyettesítését a lezserebb, gyorsabb, kevésbé gondos kiejtés okozhatja.

A magánhangzó-kapcsolatok 3,5%-nál a **hangkapcsolat helyett egyetlen hang realizálódott**. Ez lehetett a két magánhangzó közül a hosszabb időtartamú, vagy – ha mind a kettő rövid volt – esetleg valamelyiknek a hosszú párja valósult meg. Összesen 7 magánhangzó-kapcsolatnál jelentkezett: [ɛa:] → [a:], [ɔo:] → [o:], [a:ɛ] → [a:], [ɔɛ] → [ɛ], [ɛø] → [ø:], [ei] → [ɛ], [uo] → [o:]. A hangok időtartama a jelen kísérletben mért átlagokhoz képest (vö. 1. táblázat) valamivel hosszabb: [a:]: 152,9 ms, [ø:]: 159,0 ms, [ɛ]: 94,7 ms, ez alól csak az [o:] kivétel (98,8 ms). A hangkapcsolat egyetlen hanggá történő egyszerűsödésének oka egyrészt az artikulációs ökonómiában keresendő, másrészt pedig itt is a gyorsabb, kevésbé gondos kiejtés hatásait figyelhetjük meg.



10. ábra
A beszélők közötti eltérések

Az elemzések különféleképpen érvényesülő koartikulációs folyamatokat tártak fel két magánhangzó kapcsolatában. Ezzel összefüggésben felmerül a kérdés, hogy ezek a jelenségek milyen mértékben jellemzők egy adott beszélőre, illetőleg mennyire nevezhetők általános, akár nyelvspecifikus jelenségnek. A 10. ábrán hasonlítottuk össze a tíz kísérleti személytől származó adatokat. Az öt megfigyelt

koartikulációs jelenség közül csak három – a hangátmenet hiánya, a [j] hang betoldása és a hangátmenet – fordult elő minden beszélőnél, [j] hanggal történő behelyettesítést 6, míg redukálódást 4 személynél figyeltünk meg. Az ábrán látható, hogy a 10 résztvevő között igen nagy különbségek jelentkeztek, ezek a hiátustöltő [j] hangnál a legnagyobbak, ezt követi a hangátmenet hiánya, majd a rövid idejű hangátmenet.

Korrelációanalízissel további összefüggéseket tártunk fel az egyes kísérleti személyek artikulációs tempója és a beszédükben produkált koartikulációs jelenségek között. A gyorsabb tempóval beszélőknél valamivel gyakoribb a hangátmenet hiánya ($r = 0,296$, $p < 0,406$), a hangkapcsolat redukálódása egy hangra ($r = 0,270$, $p < 0,45$), illetve a [j]-vel való behelyettesítés megjelenése ($r = 0,444$, $p < 0,199$). A hiátustöltő [j] hangok nagyobb arányú használata inkább a lassúbb tempóval beszélőkre jellemző (itt közepes, negatív korrelációt találtunk: $r = -0,420$, $p < 0,0227$). A rövid idejű hangátmeneteknél nem volt összefüggés a beszédsebesség és a gyakoriság között ($r = -0,090$, $p < 0,789$). Annak ellenére, hogy a korrelációk csak közepesek, és egyik esetben sem szignifikánsak, úgy tűnik, a relatíve gyorsabb tempó (igazán gyorsan beszélő személy nem vett részt a kísérletben) a hangidőtartamok lerövidüléséhez, illetve a hangkapcsolat egyszerűsödéséhez vezethet.

Következtetések

Az elvégzett kísérlettel három fő kérdésre kerestük a választ a magánhangzó-kapcsolatokra vonatkozóan: elemeztük a kapcsolódásokban részt vevő magánhangzók formáns- és időszerkezetét; vizsgáltuk a magánhangzó-kapcsolatokban végbemenő koartikulációs folyamatok működését, illetve a beszédsebesség hatását, emellett objektív adatokkal árnyaltuk az eddigi elméleti kutatások állításait.

Megállapítottuk, hogy a kísérletben elemzett magánhangzók egy magánhangzó-kapcsolat tagjaiként is megőrzik jellemző formánsszerkezetüket (viszonyítási alapként a szakirodalmi átlagadatokat vettük). Az időtartam tekintetében fennmaradt a rövid-hosszú párok eltérése (bizonyos átfedésekkel), illetve a hangok időtartamának határértékeit vizsgálva kiderült, hogy a magánhangzó-kapcsolódás hangjainak alsó határértékei rövidebbek a szakirodalomban megadott időtartamok ugyanazon értékeinél.

Két magánhangzó között összesen öt különböző koartikulációs jelenséget adatlottunk. Leggyakoribb a hangátmenet hiánya volt (45,6%), ezt követte a [j] hang betoldása (34,6%), a hangátmenet megjelenése (13,2%), az egész hangkapcsolat redukálódása egyetlen hangra (3,5%), majd az adott magánhangzó [j]-vel történő helyettesítése (3,1%). A kapott adatok nagyobb mértékű átfedéseket mutattak a magánhangzók minősége és a közöttük megvalósuló koartikulációs jelenségek előfordulási arányai között. Például [i] hang környezetében is regisztráltunk hangátmenet nélküli kapcsolódásokat, illetve labiális és alsó nyelvvállású hangok (leginkább az [y, y:] mellett is előfordult [j] hang betoldása, holott az elméleti kutatások alapján ez nem volt várható.

A magánhangzó-kapcsolatokban regisztrált koartikulációs jelenségek előfordulási arányait befolyásolta még az adott beszélő artikulációs sebessége és egyéni „habitusa” is. A gyorsabban beszélőknél többször fordult elő a hangátmenet nélküli kapcsolódás, a behelyettesítés [j]-vel, illetve a redukálódás, míg a lassúbb tempóban beszélőkre a [j] betoldása volt a jellemzőbb.

A kapott eredmények túlnyomórészt összecsengenek a korábbi fonológiai kutatások elméleti megállapításaival. Első ízben történt meg azonban a magánhangzó-kapcsolatok akusztikai fonetikai vizsgálata, amely által objektív adatokat kaptunk az eddig csak elméleti szinten vizsgált hangjelenségekről. Az itt közölt eredmények mondatkorpuszon, igényes köznyelvi ejtésben mutatták be a két magánhangzó között létrejövő koartikulációs jelenségeket; ezek alapján válik lehetségessé az, hogy spontán beszédre vonatkozó kutatással, illetve percpációs kísérlettel a hallgató szempontjából is ellenőrizzük e vizsgálat megállapításait.

Irodalom

- Elekfi László 1992. *A magyar hangkapcsolódások fonetikai és fonológiai szabályai*. Linguistica, Series A, Studia et Dissertationes 20. MTA Nyelvtudományi Intézet, Budapest.
- Gósy Mária 1995. Szükséges és szükségtelen hangátmenetek. *Beszédkutatás* '95. 20–32.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Kassai Ilona 1999. Felnőtt nyelvhasználók szótagszerkezeti preferenciái. In Kassai Ilona (szerk.): *Szótagfogalom – szótagrealizációk*. MTA Nyelvtudományi Intézet, Budapest, 143–152.
- Lehiste, Ilse – Peterson, Gordon E. 1992. Transitions, glides, and diphthongs. In Baken, Ronald J. – Daniloff, Raymond G. (eds.): *Readings in clinical spectrography of speech*. Singular Publishing Group Inc.–Kay Elemetrics Group, San Diego, 286–295.
- Menyhárt Krisztina 2003. Zöngésedési és zöngétlenedési folyamatok a /j/-fonéma hangrealizációiban. *Beszédkutatás* 2003. 75–89.
- Olaszy Gábor 1994. Hangidőtartam-módosító kísérletek a gépi beszéd ritmusának javítására. *Beszédkutatás* '94. 140–152.
- Olaszy Gábor 2001. *A beszéd akusztikai-fonetikai elemzése és modellezése különös tekintettel a korszerű beszédépítés követelményeire*. Akadémiai doktori értekezés. Budapest.
- Siptár Péter 2002a. Hiátus. In Hunyadi László (szerk.): *Kísérleti fonetika – laboratóriumi fonológia 2002*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 85–99.
- Siptár Péter 2002b. Optimális hiátustöltés. *Beszédkutatás* 2002. 70–83.
- Vértes O. András 1982. A magyar beszédhangok akusztikai elemzésének kérdései. In Bolla Kálmán (szerk.): *Fejezetek a magyar leíró hangtanból*. Akadémiai Kiadó, Budapest, 71–115.

NONVERBÁLIS VOKÁLIS JELEK A TÁRSALGÁSBAN

Markó Alexandra

Bevezetés

A dialogikus spontán beszéd fonetikai vizsgálata közben sok olyan vokális jellel vagy jelszerű képződménnyel találkozunk, amelyek funkcióit korábban nem vagy alig vizsgálták a magyar nyelvtudományban. Ilyen az általunk „hümmögés” gyűjtőnévvel illetett hangjelenség is, melynek mind formai, mind funkcionális jellemzői sokszínűséget mutatnak. A hümmögés jelenségek első fonetikai jellegű definícióját Vértes O. Andrásnál olvashatjuk (1987), aki *hmgetés* néven tárgyalja a nyelven kívüli hangjelek között – a köhögés, a tüsszentés, a torokköszörülés, a sóhajlás, a lihegés, a nevetés és a sírás mellett. Vértes a hümmögést létrehozó artikulációs folyamatot így határozza meg: „hmgetéskor zöngét és [h]-féle hangot képzünk, ezek levegője orrunkon át távozik” (56).

Két korábbi tanulmányban (Markó 2005 és 2006) a hümmögésnek a verbális interakcióban használatos néhány formáját vizsgáltuk akusztikai és percepcióss fonetikai szempontból: az ’igen’ jelentésű és az odafigyelést jelző (amely szerkezetében egyezik az előbbivel, az ismétlődés azonban sokkal jellemzőbb rá), a ’nem’ jelentésű, valamint a kérdő funkciójú formákat (az utóbbiból semleges és emocionális változatokat) (Markó 2005); illetőleg a tiltakozó és a kérlelő hümmögést (Markó 2006). Megállapítottuk a lehetséges akusztikai kulcsokat, amelyek alapján az egyes formák eltérnek egymástól – hiszen ezek alapján dekodolhatók, s tölthetők be szerepüket –, valamint magyar anyanyelvű, illetve idegen ajkú beszélők percepcióss tesztekben adott válaszaiból következtettünk a formák funkcionális egyértelműségére és nyelvfüggő/nyelvfüggetlen voltára. Az eredmények alapján elmondható volt, hogy az elemzett három alaptípus (’igen’, ’nem’, kérdés) tagoltságában és a dallammenet jellemzőiben tér el egymástól. Az ’igen’ jelentésű hümmögéstől a figyelemhümmögés mindössze a jellemzőbb ismétlés tényében különbözik, míg a semleges kérdőtől az emocionális kérdő a hangköz értékében, vagyis a frekvenciaváltozás meredekségében tér el leginkább. Megállapítottuk, hogy a zörejelem, mely Vértes O. András definíciójában kötelező, csak egyes típusokra nézvést kötelező vagy lehetséges, a ’nem’ jelentésének kifejezésében például mintha kifejezetten kerülendő lenne. A tagolódo hümmögések (’igen’, figyelemjelzés és ’nem’) mindegyike időben átlagosan az első harmadnál oszlik két részre. A tagoltság megvalósulhat pusztán dallamívtörésben vagy a zöngé időszakos kimaradásában, és típustól függően ezek kombinálódhatnak a zörejelem megjelenésével is. Ha a zöngé nem szakad meg a zörejelem megjelenésekor, mint az ’igen’ típusú minták némelyikénél, akkor ke-

vert jel jön létre. Az alaptípusok nem teljes mértékben nyelvfüggők, legalábbis a magyaron kívül sok más európai nyelv is hasonlóan él velük – ezt bizonyították az általunk elvégzett percepciók tesztek.

A tiltakozó és a kérlelő minták nem mutatnak tagoltságot, dallamformájukban hasonlóan egymáshoz, de egyértelműen eltérnek a korábban vizsgált hümmögésektől. A hangmenet feltétlenül akusztikai kulcsként funkcionál például a kérdő formától való elkülönítésben. A tiltakozás és a kérlelés elsődleges elhatárolója a formánsstruktúra (ha ez a fokozott emocionalitás miatt háttérbe szorul, a két funkció is összeolvad a parancs aktusába), másodlagos kulcsként pedig a kérlelő formák hosszabb időtartama is segíti a feldolgozást.

A korábbi eredményeket úgynevezett elicitált, vagyis **kísérletileg előhívott** mintákból álló korpusz vizsgálatából nyertük, amelyet tehát úgy rögzítettünk, hogy a beszélőknek (10-10 magyar anyanyelvű nő és férfi) megadott kommunikációs tartalmakat kellett kifejezniük ezzel a (hang)eszközzel: egy feladatlapon megadott instrukciók alapján kellett hümmögniük igenlően, kérve, nemlegesen, odafigyelést jelezve, hitetlenkedően kérve, elismerően, kérlelően és tiltakozva. A felvételeket átlagosan egyórás időkülönbséggel megismételtük, így összesen típusonként 40 hümmögés szolgálhatott a vizsgálat alapjául.

A jelen tanulmányban valódi **társalgási szituációban** rögzített hümmögések fonetikai vizsgálatát végezzük el. A kapott akusztikai adatok elemzése ezúttal is kiegészül a percepciók tesztelés eredményeinek bemutatásával, illetőleg végül összehasonlítjuk a korábbi, elicitált minták és a társalgás hümmögéseinek jellemzőit. Így választ kaphatunk arra a – kísérleti tudományokban (így a fonetikában is) gyakran felmerülő – kérdésre, hogy a laboratóriumi körülmények között zajló vizsgálatok mennyiben képesek tükrözni a mindennapi élet spontán (beszéd)jelenségeit.

Kísérleti személyek, anyag és módszer

Az alábbiakban elemzett társalgást négy 21 éves egyetemi hallgatóval rögzítettük. A kísérleti személyek mindannyian budapesti születésűek, a köznyelvi standardot beszélik, nem éltek hosszabb ideig sem magyar nyelvjárási területen, sem külföldön. Monolingvis családból származnak, de mindannyian jól beszélnek két vagy több idegen nyelven. Nincs sem beszédhibájuk, sem ismert hallásproblémájuk. Adatközlőink a társalgáshoz nem kaptak instrukciókat, mindössze az volt a feladatuk, hogy beszélgessenek, s mivel egymást jól ismerő személyek voltak, hamar természetessé vált a mesterségesen teremtetett szituáció. A beszélgetést egy professzionális rádióstúdióinak használt csendesített szobában vettük fel, térmikrofonnal, és közvetlenül számítógépre rögzítettük digitális formában. A felvétel időtartama közel két óra (113,41 perc, azaz kb. 1 óra 53 perc).

A társalgásban összesen 63 hümmögést annotáltunk, körülbelül egyenlő arányban a négy adatközlőtől. Ezek közül 30 nem elemezhető, mert valamely beszédpartner közlése alatt hangzik el, további 3-at pedig azért nem elemeztünk, mert funkciójuk erősen szituáció- és kontextusfüggő, a felvétel alapján nem

dönthető el, hogy mi lehetett pontosan a közlési szándék. Valószínűsítjük, hogy ezek a nem vokális nonverbális jegyekkel, elsősorban az arcjátékkal együtt nyelik el jelentésüket. Vizsgálatunk anyagát tehát végül összesen 30 hümmögés képezte, ezek többsége, 22 darab, férfi adatközlőtől származik, és csak 8 esetben nő a beszélő. Funkciójuk szerint ezek elsősorban igenlő-figyelemjelző hümmögések (22 előfordulás, a minták nemek szerinti aránya a férfiak javára 17 : 5); 7 hümmögés kérdő jellegű (ezek közül 2 származik női adatközlőtől); és mindössze egyetlen nemleges hümmögés van ebben a korpuszban (ezt női beszélőtől rögzítettük). Megjegyezzük, hogy az összes annotált hümmögés túlnyomó többsége is igenlő, jóval ritkábbak a kérdőek, és nem találtunk több nemleges funkciójú hangjelenséget az egész társalgásban sem.

Az akusztikai elemzést a Praat 4.3 verziószámú jelfeldolgozó szoftverrel végeztük, a rögzített hangjelenségek idő- és dallamszerkezeti jellemzőit vizsgáltuk.

A percepciók kísérlethez hét mintából álló tesztanyagot állítottunk elő, amelyben minden adatközlőtől minden nála előforduló típusra volt egy-egy példa. Így négy igenlő, két kérdő és egy nemleges hümmögés hangzott el az összeállított felvételen, véletlenszerű sorrendben a női és a férfi adatközlőktől. A tesztanyagot kétszer játszottuk le a hallgatóknak: az első meghallgatás után azt kértük tőlük, hogy állapítsák meg, „mit jelent” az adott minta, a második fordulóban pedig azt, hogy az általunk megadott jelentések (*Igen. Nem. Mit mondtál?*) valamelyikéhez sorolják be az illető sorszámú mintát. Természetesen itt választhaták azt a lehetőséget is, hogy a hümmögés egyik „jelentésnek” sem felel meg. Ezt a második értékelést kontrollként használtuk annak érdekében, hogy a különböző tesztelők által megadott jelentések szinonimitása ne befolyásolja szubjektíven a kapott válaszok feldolgozását.

A percepciók tesztben 23 magyar és 23 nem magyar anyanyelvű egyetemista és egyetemistát végzett személy vett részt. A magyar tesztelők közül 15 nő és 8 férfi, átlagéletkoruk 32 év (23 és 62 év közöttiek), mindannyian Budapesten vagy az agglomerációban élnek. A nem magyar anyanyelvűek között 1 belga (flamand) nő, 3 finn (1 nő és 2 férfi), 1 holland férfi, 10 német (5 nő és 5 férfi), 3 olasz (2 nő és 1 férfi), 1 osztrák nő, 2 spanyol (1 nő és 1 férfi), 1 svéd nő és 1 török nő vett részt a kísérletben, összesen tehát 13 nő és 10 férfi. Átlagosan 25 évesek (az életkoruk 19 és 71 év között szóródott). Nekik a kísérlet idején még csak minimális magyar nyelvi tapasztalatuk volt, éppen elkezdtek magyarul tanulni, de alig kommunikáltak ezen a nyelven. Az instrukciókat egységesen angolul kapták, és angolul vagy németül válaszolhattak.

Eredmények

A társalgásbeli **igenlő** hümmögések kb. egy szótagnyi időtartamban realizálódnak, a nők mintái valamivel hosszabbak, mint a férfiakéi (1–3. táblázat). Idő szerkezetük kb. egyharmad-kétharmad arányú egységeket tartalmaz. A 22 mintából egy az alaphanggörbe megszakadásával realizálódott, a többi folytonos, mindegyikre jellemző azonban a tagoló pont környékén a [h]-szerű zörejelem

megjelenése. A többség ezen szakaszában tehát kevert gerjesztésű a jel. A minták dallammenete relative lapos rajzolatú, de három kivétellel mindegyikre jellemző az ereszkedő-emelkedő hangmenet. Az 1. a) ábrán egy női és egy férfi adatközlő igenlő hümmögésének dallamgörbéje látható, a vonalak szinte párhuzamosan futnak. A hangterjedelem szűk minden adatközlő – még a nők – esetében is. Az alaphérvencia-adatokat ugyancsak az 1–3. táblázat összegzi. A három kivételként említett dallammenetű realizáció időbeli és hérvenciajellemzői semmiben nem térnek el a többi igenlő hümmögéstől, azonban a dallamzár itt nem emelkedő, hanem az első részt folytató, ereszkedő rajzolatú – vö. 1. b) ábra. Ezeknek a jelentésárnyalata is más: az ereszkedő zárlat miatt inkább lezáró, beleegyező, elfogadó, ráhagyó jellegűek – szemben az emelkedő végű pozitívabb attitűdöt, folytatásra felszólítást sugalló hümmögésekkel.

1. táblázat: Az igenlő minták összesített számszerű eredményei

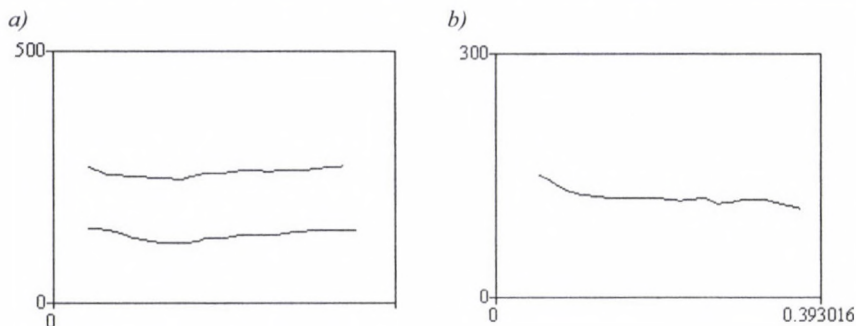
| | Időtartam (ms) | F ₀ -minimum (Hz) | F ₀ -maximum (Hz) | Hangköz (F ₀ -max/F ₀ -min) |
|------------------|-------------------|---------------------------------|---------------------------------|--|
| Alsó határérték | 167 | 112 | 132 | 1,08 |
| Felső határérték | 373 | 245 | 292 | 1,58 |
| Átlag | 280 | 151,18 | 187,70 | 1,24 |
| Átlagos eltérés | 0,04 | 31,81 | 43,37 | 0,10 |

2. táblázat: A nők igenlő mintáinak összesített számszerű eredményei

| | Időtartam (ms) | F ₀ -minimum (Hz) | F ₀ -maximum (Hz) | Hangköz (F ₀ -max/F ₀ -min) |
|------------------|-------------------|---------------------------------|---------------------------------|--|
| Alsó határérték | 263 | 202 | 250 | 1,11 |
| Felső határérték | 373 | 245 | 292 | 1,34 |
| Átlag | 321 | 220,60 | 272,20 | 1,24 |
| Átlagos eltérés | 0,02 | 11,12 | 13,44 | 0,07 |

3. táblázat: A férfiak igenlő mintáinak összesített számszerű eredményei

| | Időtartam (ms) | F ₀ -minimum (Hz) | F ₀ -maximum (Hz) | Hangköz (F ₀ -max/F ₀ -min) |
|------------------|-------------------|---------------------------------|---------------------------------|--|
| Alsó határérték | 167 | 112 | 132 | 1,08 |
| Felső határérték | 346 | 154 | 230 | 1,58 |
| Átlag | 268,00 | 130,77 | 162,12 | 1,24 |
| Átlagos eltérés | 0,04 | 8,57 | 18,15 | 0,11 |



1. ábra

Az igenlő minták jellemző dallammenetei:

a) ereszkedő-emelkedő; b) ereszkedő

(Az x tengelyen az idő, az y tengelyen a frekvencia látható.)

A percepció tesztben csak a tipikusabbnak mondható ereszkedő-emelkedő dallammenetű minták szerepeltek: minden adatközlőtől egy-egy, tehát összesen négy. A magyar tesztelők a különböző igenlő hümmögéseket az első fordulóban 82,6–95,7%-ban ismerték fel, ez az összes lejátszott mintát tekintve 89,1%-os átlagot jelent. A kontrolltesztben ezek az értékek 69,6 és 100% között adódtak, ez átlagosan 90,2%-os felismerés.

A nem magyar anyanyelvűek 52,2–82,6%-ban, átlagosan 63,0%-ban tartották igenlőnek a mintákat az első tesztfordulóban, a kontrolltesztben a szélső értékek ugyanezek, de az átlag valamivel magasabb: 68,5%. A magyar tesztelők számára a magyar beszélők igenlő hümmögése tehát sikerebben volt dekódolható.

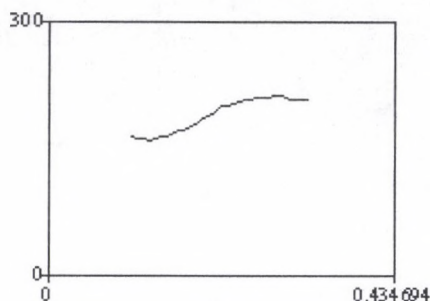
A tesztelők által megadott jelentések közül az első fordulóban olyanokat fogadtunk el az igenlő (illetve a formailag ezzel azonos figyelemjelző) hümmögésre, mint *egyvetértés*, *figyelem*, *helyeslés*, *kapcsolattartás*, *tudomásulvétel*, *jóváhagyás* stb.; illetve *Értem*, *Persze*, *Jó*, *Szerintem is*, *Egyvetértek*, *Valószínűleg*, *Aha*.

A hét **kérdő** hümmögés mindössze két adatközlőtől származik: az egyik férfi adatközlőtől öt, az egyik nőtől pedig kettő. A számadatokat a 4. táblázat összegzi – a kevés adat és adatközlő miatt nem különítettük el őket nemek szerint.

4. táblázat: A kérdő minták összesített számszerű eredményei

| | Időtartam (ms) | F ₀ -minimum (Hz) | F ₀ -maximum (Hz) | Hangköz (F ₀ -max/F ₀ -min) |
|------------------|-------------------|---------------------------------|---------------------------------|--|
| Alsó határérték | 165 | 157 | 190 | 1,17 |
| Felső határérték | 221 | 265 | 326 | 1,32 |
| Átlag | 201,00 | 192,86 | 240,00 | 1,25 |
| Átlagos eltérés | 0,02 | 38,65 | 48,29 | 0,04 |

A kérdő hümmögések időtartama valamivel rövidebb, mint az igenlő hümmögéseké, nem tagolódnak, és nem jelenik meg bennük [h]-szerű spirantikus elem. A dallam jellemzően emelkedő, a hangköz nem tágabb, mint az imént vizsgált típusban, de magasabb hangfekvésben realizálódik (lásd 2. ábra).



2. ábra

A kérdő hümmögések jellemző dallammenete

(Az x tengelyen az idő, az y tengelyen a frekvencia látható.)

A magyar tesztelők a két lejátszott hümmögést 87,0 és 91,3%-ban értékelték kérdő jellegűnek mind az első, mind a második tesztfordulóban, ez átlagban 89,1%-ot jelent.

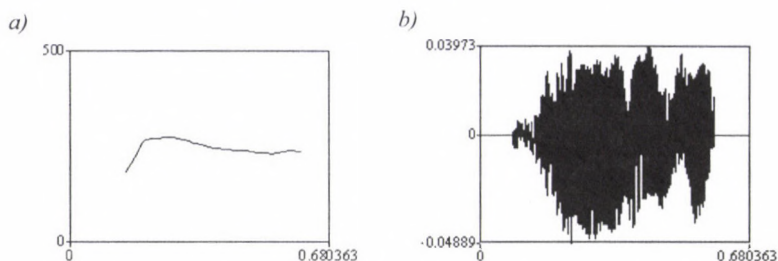
A nem magyar anyanyelvűek 69,6 és 82,6%-a találta kérdő jellegűnek a mintákat az első tesztfordulóban, átlagosan tehát 76,1%-uk. A kontrollteszt 69,6 és 78,3%-os eredményt hozott a két mintára, összességében tehát 73,9%-os volt a felismerés. A magyar anyanyelvűek a kérdő hümmögést ugyancsak valamivel sikeresebben dekódolták.

Az itt megadott jelentések közül kérdőnek fogadtuk el a kérdő mondatok (*Tessék? Mit mondtál?* stb.) mellett a *visszakérdezés*, *érdeklődés* stb. válaszokat.

Az egyik női adatközlő egyetlen vizsgálható **tagadó** hümmögése hosszabb a korábbiaknál (5. táblázat). A hangköze is tágabb, de egyetlen adat alapján ez nem tekinthető szabályszerűnek. A dallammenete szökő indítással dominánsan ereszkedő – vö. 3. a) ábra. Hallhatóan tagolódik körülbelül a közepén, de nem F_0 -változás tagolja (az ereszkedés folyamatos), hanem intenzitásminimum – vö. a 3. b) ábrán látható rezgésképpel.

5. táblázat: A nemleges minta eredménye

| Időtartam (ms) | F_0 -minimum (Hz) | F_0 -maximum (Hz) | Hangköz (F_0 -max/ F_0 -min) |
|-------------------|------------------------|------------------------|--------------------------------------|
| 389 | 181 | 275 | 1,52 |



3. ábra. A tagadó hümmögés: a) dallammenete, b) rezgésképe
(Az x tengelyen az idő; az a) ábra y tengelyén a frekvencia,
a b) ábra y tengelyén az amplitúdó látható.)

A nemleges hümmögés 82,6 és 87,0%-os eredményt hozott az első és a második fordulóban a magyarok esetében. A nem magyar anyanyelvűek az első tesztben 56,5%, a kontrolltesztben viszont már 87,0%-ban találták tagadó jelentésűnek a mintát, itt tehát azonos a magyar és az idegen ajkú tesztcsoport eredménye.

A kapott válaszok közül a nemlegeshez soroltuk az alábbiakat: *tiltakozás, ellenkezés, egyet nem értés, tagadás* stb., illetve a *Nem hiszek neked. Hagyd abba! Nincs kedvem.* típusú mondatokat.

A célzott kísérletben előhívott és a valódi társalgásban megjelent hümmögések akusztikai és perceptív vizsgálatát tehát azonos módszertani feltételek között végeztük el. A két elemzés eredményei közül azonban a minták elemszáma miatt csak az igenlő és a kérdő típusú hümmögésről tett megállapításainkat vehetjük össze reprezentatív módon. Gyakoriságuk miatt ugyanakkor ezek a típusok is tűnnek a legjelentősebbeknek, ami összefügghet azzal is, hogy ezek a hümmögések, de különösen az igenlő-figyelmjelző típusú a formálishoz közelebb álló társalgási helyzetekben is gyakorta előfordul.

Az igenlő hümmögésről mindkét vizsgálat megállapította, hogy 1 : 2 arányban tagolódik, a tagoló ponton kötelezően spirantikus zörej jelenik meg, amely az esetek kisebb részében a zöngé rövid idejű megszakadásával jár együtt, általában azonban folyamatos az alaphanggörbe. Az időtartamátlagok között kb. 20%-nyi különbség van: a kísérletileg előhívott minták átlagosan ennyivel hosszabbak, ami pusztán a tervezettségre vezethető vissza. A hangközértékekben ugyanez mutatkozik meg. A dallamgörbéket összehasonlítva azt látjuk, hogy a célzott kísérletben több meredekebb mintát is rögzítettünk, míg a társalgásban felvettek relatíve lapos rajzolatúak. Ez a magyarázat tehát a hangközértékek közötti különbségre.

A kísérletben előhívott minták révén vizsgálhattuk az úgynevezett figyelmjelző hümmögéseket is, és megállapítottuk, hogy ez a típus szerkezetileg, időtartamában és frekvenciajellemzőiben azonos az igenlő hümmögéssel, az egyetlen különbséget abban érhetjük tetten, hogy az ismétlődés tényét a figyelmjelzés

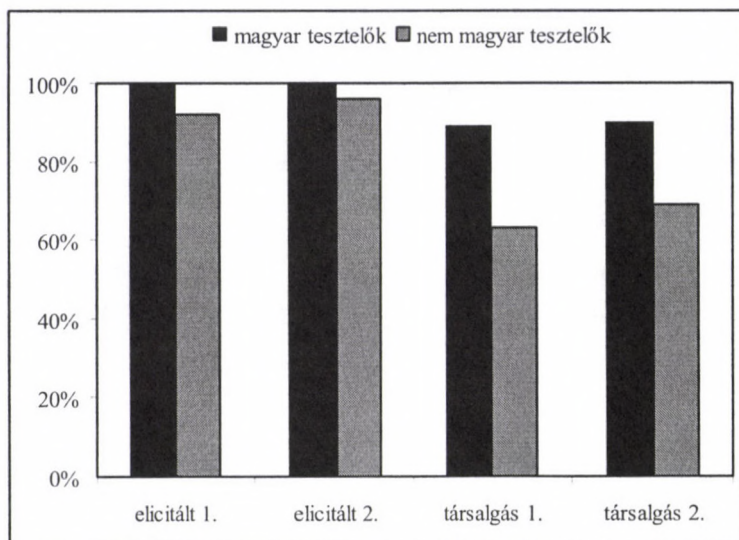
esetében a percepciók tesztelők releváns jegynek tartották. A társalgásban előforduló és elemezhető minőségű hümmögések között ilyen ismétlődést mégsem találtunk. Ennek több oka is lehet: a figyelmünkről jellemzően akkor biztosítjuk a beszédpartnerünket, amikor ő éppen beszél, így a hümmögés részben vagy egészen egyszerre hangzik a beszélő közléseivel. Az is valószínűsíthető, hogy kettőnél több szereplős társalgás esetén a többi résztvevő jelenléte miatt a partnereknek nem kell gyakran meggyőzniük figyelmükről a beszélőt. Mindemellett közvetlen kontaktus esetén a partnerek nonverbális jeleinek, gesztusainak, mimikájának, testtartásának stb. látványa is elegendő, nincs szükség feltétlenül a vokális jelekre is – szemben a telefonbeszélgetésekkel, amikor a vokális jelek hiánya meg is akaszthatja a beszélgetést. Érdekes ugyanakkor, hogy a percepciók tesztben lejátszott igenlő hümmögéseknek a tesztelők egy része kifejezetten figyelemjelző funkciót tulajdonított.

A semleges kérdő minták esetében belső tagolódást egyik korpuszban sem figyeltünk meg, a hümmögés emelkedő alaphangfrekvenciájú zöngében realizálódik. A kísérletben előhívott minták közül néhányban a zöngéindítás hehezetes volt, a spontán társalgásból előállt korpuszban ilyet nem találtunk. Az átlagidőtartam a két anyagban szinte azonos, a hangközértékek átlaga azonban a kísérletileg előhívott korpuszban ismét nagyobb mintegy 25%-kal. Ennek okát is a kísérleti helyzet jellegében kereshetjük.

Ugyan a társalgásban csak egy nemleges hümmögést rögzítettünk, erről megállapítható, hogy mind tagolódását, mind időszerkezeti, mind pedig frekvenciajellemzőit tekintve a kísérletben vizsgált két nemleges altípus közötti átmenetként realizálódott (vö. Markó 2005). Ez alapján akár az is feltételezhető, hogy a nemleges hümmögés variánsok kontinuumot alkotnak, amelynek az altípusok elemei és a spontán társalgásbeli realizáció is részei. A kísérletben előhívott minták percepciók tesztelése alapján a nemleges hümmögés prototípusa szünettel megszakított és meredekebb, szökő-eső kezdetű dallamformával rendelkezik, az ettől eltérőknél valamely kiegészítő kulcsra (pl. fejrázás, mimika) is szükség lehet a felismeréshez.

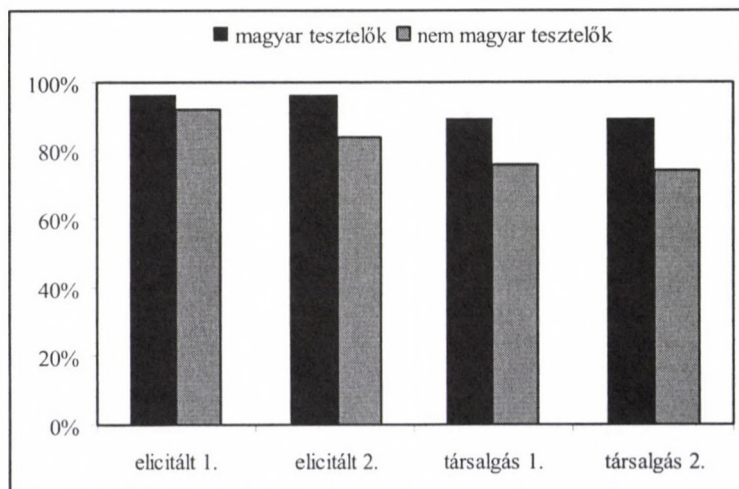
Összegezve az akusztikai elemzés eredményeit: a kísérletileg előhívott és a társalgásbeli spontán minták között lényeges – minőségi – eltérést nem találunk, a mennyiségi különbségeket a képzelt szituációbeli markírozás okozhatja, a beszélőnek azon szándéka, hogy a kifejező prototípust valósítsa meg; illetőleg az a tény, hogy a kontextus és a nem vokális nonverbális elemek hiányában a vokális jeleknek kell kifejezőbben, ezeket pótolva is megvalósulniuk.

A percepciók tesztjei eredményeit összevetve mindegyik típusban azt találjuk, hogy a társalgásbeli hümmögések dekódolása nehezebb volt, mint a kísérletileg előhívott mintáké. A 4–6. ábrán részletesen is bemutatjuk a teszteredmények eltéréseit. Látható, hogy míg a kísérletileg előhívott igenlő és nemleges hümmögések a magyarok számára 100%-ban felismerhetők voltak, és a kérdők is megközelítették ezt az értéket, a valódi társalgásból kivágott hümmögések legfeljebb 90%-os eredményt hoztak, vagyis valamivel bizonytalanabb volt a percepció.



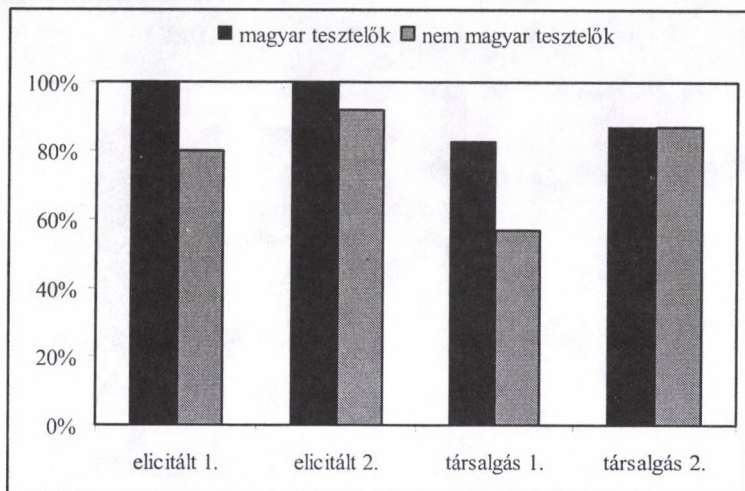
4. ábra

Az igenlő hümmögések percepciójának eredményei



5. ábra

A kérdő hümmögések percepciójának eredményei



6. ábra

A nemleges hümmögések percepciójának eredményei

A kísérletileg előhívott minták mindegyike esetében a nem magyar anyanyelvűek eredményei alacsonyabbak a magyarokénál, és bár ez a társalgásbeli hümmögések többségére is igaz, a nemleges hümmögések kontrolltesztjében mégis ugyanolyan eredményt értek el a nem magyar anyanyelvűek, mint a magyar tesztelők. Különösen érdekes, hogy ez a kísérletileg előhívott nemleges hümmögés első fordulójában elértnél is jobb eredmény, ami abból adódhat, hogy a szituációból kivágott, kontextus nélküli, és bizonyos értelemben kevésbé markáns hümmögést könnyebb volt előzetesen megadott kategóriák alapján besorolni, mint első hallásra jelentést tulajdonítani neki.

A kísérletileg előhívott mintákra kapott jobb eredmények azt mutatják, hogy az akusztikai szempontból prototipikusabb megvalósulás a percepció számára is prototípusként értelmeződik, hiszen kontextus hiányában sokkal nehezebben lehet felismerni az aktualizált, valamilyen mértékben érzelmileg színezett, változékonyabb társalgásbeli formákat. A nem magyar anyanyelvűek olykor jóval alacsonyabb felismerési aránya a társalgásbeli hümmögéseket illetően szintén azt mutatja, hogy ezek a realizációk sokkal jobban függenek az adott szituációtól, kontextustól, beszélőtől, érzelmi töltéstől, és talán az adott nyelvközösségre jellemző kommunikációs sajátosságoktól is.

Következtetések

A társalgásban annotált összes hümmögés kb. 70%-a volt igenlő típusú, további 17,5% pedig kérdő. Valószínűsítjük, hogy ezek az arányok a hétköznapi informális társalgásokban megjelenő hümmögések megoszlását is jelzik. A cél-

zott kísérletben a hűmmögések nem spontán módon realizálódtak, hanem a kísérletben részt vevők körülírások alapján hozták létre őket. Ebből fakadóan mind a típusok aránya, mind a nemek részvétele kiegyenlítettebb volt a tervezett kísérleti korpuszban. Ugyanakkor, mivel a beszédsszituáció képzeletbeli volt, a spontán realizációkkal szemben az így előhívott minták tervezettebbek, ezért valószínűleg közelebb állnak a nyelvhasználók kommunikatív kompetenciája részeként tárolt hűmmögés-prototípusokhoz. Ez alapján a kísérletben rögzített hűmmögésrealizációk bizonyos tekintetben kevésbé változatosak, mint a konkrét szituációban spontán aktualizálódó variánsaik.

A percepció tesztek bizonyítékokat szolgáltatnak arról, hogy az itt tárgyalt alaptípusok nem teljes mértékben nyelvfüggők, legalábbis a magyaron kívül sok más európai nyelv is hasonlóan él velük. Vértes O. András (1987: 56) azt írja: „A hűmmögést illetően fölmerül a kérdés, hogy funkciója és formái általánosak-e vagy sem. Finnországi vizsgálódásom szerint a finnek nem értették a magyar hűmmögő »szavak« egy részét: tehát ebben még rokon nyelvek között sem kell teljes megegyezésnek lennie.” Sajnos nem tudjuk, milyen típusú hűmmögéseket hogyan és pontosan milyen eredményekkel tesztelt a szerző, így a saját eredményeinkkel nem tudjuk összevetni azokat. Adataink alapján az azonban bizonyosnak látszik, hogy az európai nyelvek egy részében (areális alapon?) a vizsgált alaptípusok használatosak, még ha a variabilitás határai nyelvenként eltérőek is. A percepció teszteknel nem találtunk nyelvspecifikus eltéréseket, a jobban szóródó válaszoknál sem voltak felfedezhetők olyan mintázatok, amelyek anyanyelvi motiváltságra utaltak volna (ami egybevág Péteri 2006 eredményeivel), ezt azonban az azonos anyanyelvű adatközlők alacsony száma is magyarázhatja.

A társalgás és a célzott kísérlet korpuszának hűmmögései jellegzetes „mennynységi” különbségeket mutatnak: a kísérleti anyagban markírozottabbak a realizációk, mint a valódi spontán helyzetben. Ezt a tanulságot kiterjesztve feltételezzük, hogy más nyelvi adatok spontán és „laboratóriumi” realizációit összevetve ugyancsak jelentős különbségeket találhatunk a kidolgozottság, a tervezettség, az előkészítettség tekintetében. Vagyis elképzelhető, hogy a laboratóriumi korpuszok olyan ideális formákat tartalmaznak, amelyeket a normál beszédtevékenység során nem vagy alig tapasztalunk, csak prototípusokként írhatók le. Ugyanakkor a laboratóriumi korpuszok sokkal alkalmasabbak arra, hogy sok adatközlőtől nyerjünk nagyszámú vizsgálható hangzó jelenséget, így statisztikai relevanciájuk is van – szemben a természetes szituációkban rögzített beszédkorpuszokkal, amelyeknél az adatközlők és a vizsgálandó realizációk száma csak igen hosszadalmas gyűjtő és elemző munka révén érheti el a statisztikai elemzésre is alkalmas, számottevő mennyiséget. A legertermészetesebb beszédsszituáció, a társalgás elemzésében különösen gyakori problémát jelent, hogy az egyszerre beszélés vagy más „zajok” (nevetés, köhögés vagy éppen hűmmögés), azaz a többi adatközlő aktív beszélőként való jelenléte korlátozza az akusztikai elemzésre alkalmas közlések mennyiségét, tehát sokkal nagyobb időtartamú korpuszok vizsgálata hozhat csak hasonló mértékben általánosítható eredmé-

nyeket, mint a monologikus beszédprodukciók vagy a még kevésbé természetes laboratóriumi beszédprodukciók elemzése. Többek között ez is alátámasztja a spontán beszéd egyre szélesebb körű és mélyebb vizsgálatának fontosságát.

Irodalom

- Markó Alexandra 2005. „Szavak nélkül”. Nonverbális vokális közlések fonetikai elemzése. *Magyar Nyelvőr* 129. 88–104.
- Markó Alexandra 2006. A hümmögés mint beszédaktus. In Mártonfi Attila – Papp Kornélia – Slíz Mariann (szerk.): *101 írás Pusztai Ferenc tiszteletére*. Argumentum, Budapest, 604–612.
- Péteri Attila 2006. Prozódia a grammatika és a pragmatika határán. Egy német–magyar összehasonlító vizsgálat eredményei alapján. *Beszédkutatás* 2006. 69–85.
- Vértes O. András 1987. *Bevezetés a magyar hangstilisztikába*. Nyelvtudományi Értekezések 124. Akadémiai Kiadó, Budapest.

A tanulmány az OTKA T046455 számú pályázatának a támogatásával jött létre.

PROZÓDIA A GRAMMATIKA ÉS A PRAGMATIKA HATÁRÁN

Egy német–magyar összehasonlító vizsgálat eredményei alapján

Péteri Attila

Bevezetés

A germanisztikai szakirodalomban többen utalnak arra, hogy a mondatfajták, illetve a mondattípusok elhatárolásának egyik alapvető kritériuma az intonációs jelöltség, amely elsősorban egyes kérdő és felszólító mondattípusokat különböztet meg az azonos grammatikai szerkezettel rendelkező kijelentő mondatoktól. Más, mondattipológiával és mondatmodalitással foglalkozó kutatók viszont explicite tagadják a mondatintonáció grammatikai szerepét, és amellet érvelnek, hogy az intonációs és prozódiai eszközök a pragmatika szintjén fejtik ki hatásukat, vagyis a megnyilatkozás pragmatikai funkcióit, illokúciós tartalmát fejezik ki.

A szakirodalomban fellelhető igen különböző nézetek tanulmányozása során fordult figyelmem a következő elméleti kérdés felé: A mondat intonációs és prozódiai jegyei mely szinten fejtik ki hatásukat, milyen viszonyban állnak a mondat grammatikai típusával, a beszélő attitűdjeivel, illetve illokúciós szándékával.

A német nyelvészetben a mondatintonáció vizsgálata egészen a 90-es évekig alapvetően szintaktikai orientáltságú. Otto von Essen 1964-ben megjelent, a modern germanisztikai intonációkutatás nyitányát jelentő monográfiájában az intonáció elsődleges szerepét a mondattípusok jelölésében, nevezetesen a kijelentő és az eldöntendő kérdő mondat elkülönítésében, valamint az összetett mondatok tagmondatainak elhatárolásában látja. A későbbi differenciáltabb vizsgálatok is, mint Isačenko–Schädlich (1971), Pheby (1980), Altmann (1987, 1993), bár az intonáció szerepét nem szűkítik le csupán a mondattípus jelölésére, mégis az intonációnak kizárólag vagy elsődlegesen grammatikai szerepet tulajdonítanak. Ehhez hasonlóan a *Magyar grammatika* (Keszler szerk. 2000: 379 kk.) mondattipológiájában is az intonáció a mondattípusok elkülönítésének egyik prominens eszköze. Wunderlich (1988), aki az autoszegmentális fonológia keretében vizsgálja a német mondat dallamát, a mondatvégi intonáció szempontjából szintén grammatikai jelentéssel bíró tonális morfémaokról beszél.

A 90-es évek elején az elsősorban pragmatikai orientáltságú angolszász intonációkutatás hatására a germanisztikában is egyre erőteljesebben fogalmazódik meg a gondolat, hogy a mondatintonáció a mondat grammatikai szerkezetétől többé-kevésbé függetlenül elsősorban a megnyilatkozás pragmatikai funkcióját jelöli. Így véli például a Cruttenden (1977) modelljét alkalmazó Kohler (1991),

és a Pierrehumbert (1980) nyomán autoszegmentális keretben dolgozó Féry (1993) is. Brandt és munkatársai (1992) amellett érvelnek, hogy az emelkedő mondatvégi intonációval realizált kijelentő mondat az adott szituációban kérdő illokúciós funkciót kap anélkül, hogy grammatikai szerkezete és ezáltal a mondat típusa megváltozna.

A német intonáció- és proszódia kutatásban mérföldkönek számít Margret Selting 1995-ben megjelent *Prosodie im Gespräch* című monográfiája, amely szerint a beszéd valamennyi proszódiai sajátosságát, így az intonációt is, a grammatikától függetlenül csak az etnometodológiai konverzációelemzés keretében lehet adekvát módon leírni. Selting egy több kísérleti személy részvételével hangstúdióban készített kétórás informális beszélgetésből vett nyelvi példákkal próbálja bizonyítani, hogy a német nyelv valamennyi intonációs mintája kombinálható különböző grammatikai szerkezetű mondatokkal. Így szerinte mind az eldöntendő, mind a kérdőszavas kérdő mondat záródhat eső és emelkedő intonációval is, a mondat végi beszéddallam pedig azt jelöli, hogy pragmatikailag zárt, vagyis tömör, lényegre törő választ igénylő, vagy pedig pragmatikailag nyitott, a konverzációt tovább vezető, probléma-, illetve témafelvető kérdéssről van-e szó. Rabanus (2001) Selting modellje alapján a német és az olasz nyelv intonációs mintáit összevetve olyan nagymértékű hasonlóságot vél felfedezni, amelynek alapján az intonációt az egyes nyelvtől független jelrendszernek tekinti. Igaz ugyan, hogy Rabanus a vizsgált példákat meglehetősen egyoldalúan választja ki, legtöbbször a kiemelés-hangsúlyozás különböző variációit demonstrálja.

Selting modelljét egyoldalúsága miatt bírálja Hetland és Molnár (1995), akik részben azt vetik Selting szemére, hogy az általa hivatkozott szakirodalmat néha leegyszerűsítve interpretálja, és nem veszi észre az abban képviselt differenciált felfogásokat, részben pedig azt, hogy mivel a Selting korpuszában szereplő szövegpéldák nem képeznek egymással szisztematikus oppozíciót, ezért a vizsgálatból a vizsgálati módszer következményeként csak az intonáció pragmatikai funkciói válnak megfigyelhetővé, míg az esetleges grammatikaiak rejtve maradnak.

A magyar intonációkutatás szakirodalmát áttanulmányozva úgy tűnik, hogy az intonáció grammatikai vagy pragmatikai szerepével kapcsolatos kérdés kevésbé jelenik meg explicit módon, bár a hagyományos módszerekkel dolgozó kutatók (pl. Elekfi 1962, Fónagy–Magdics 1967, Fónagy 1998) inkább grammatikai, az autoszegmentális fonológia keretében gondolkodók (pl. Varga 1998, 2001, 2003) inkább pragmatikai funkciókat tulajdonítanak a leírt intonációs típusoknak. Nagyobb hangsúlyt kapnak azonban a magyar intonáció- és proszódia kutatásban a kísérleti fonetikai vizsgálatok (Gósy–Terken 1994; Gósy 2004; Olaszky 2001, 2002), amelyekben a kutatók a kérdés explicit módon történő felvetése nélkül az intonációs és proszódiai eszközöknek mind grammatikai, mind pragmatikai funkcióit leírják.

Anyag, módszer, kísérleti személyek

A német és a magyar nyelv intonációs és prozódiai mintáinak kontrasztív összevetése aktuális, sőt sürgető, hiszen mindkét nyelv oldalán számos tudományos leírást találunk, miközben összehasonlító tanulmány tudtommal mindeddig még nem készült. A Praat (verziószám: 4.3) elemző program segítségével 11 német és 9 magyar kísérleti személy hangmintáit elemeztem és hasonlítottam össze egymással 2005 tavaszán és nyarán. Kérdésfelvetéseim alapvetően grammatikai szempontúak voltak. Bár az alább leírt vizsgálat fő célja nem a német és a magyar nyelv prozódiai rendszerének fonetikai összehasonlítása, hanem a feltett elméleti alapkérdés megválaszolása a kutatások továbbvitele céljából, az általam végzett vizsgálat szükségszerűen érinti a két nyelv prozódijának számos hasonlóságát és különbségét.

Vizsgálatom a német és a magyar nyelvészeti szakirodalomból indul ki, az empirikus megfigyelések német és magyar nyelvi példákra épülnek. Reményeim szerint azonban a konkrét részeredmények magukban rejtik bizonyos generalizált következtetések lehetőségét is (legalább a hipotézis szintjén). Azt remélem, hogy egy ilyen, a nyelvtudomány egyes részterületei határát érintő vizsgálat felhasználható adalékkul szolgálhat mind a fonetika, mind pedig a grammatika művelőinek, és értékes lehet egy holisztikus alapon álló integratív nyelvfelfogás szempontjából is.








A kísérleti személyeket arra kértem, hogy általam írt dialógusrészleteket minél természetesebb kiejtéssel olvassanak fel. A magyar beszélők az ELTE magyar anyanyelvű német szakos hallgatói voltak, a németek pedig Braunschweigenben élő középiskolás és egyetemista diákok, valamint két idősebb, 40-45 éves férfi. A német résztvevők valamennyien a standard német nyelv beszélőinek tartották magukat, nem pedig valamelyik dialektushoz tartozónak (ahogy ők kifejezték: „Ich habe keinen heimatlichen Dialekt” vagy „Ich spreche normales Hochdeutsch”). A beszélők a dialógusrészleteket egyszer olvasták át annak érdekében, hogy a felolvasás során ne akadozzanak, de nem gyakorolhatták. Ezzel próbáltam elkerülni a mesterkélt felolvasási módokat. A hangminták kiértékelésekor figyelembe vettem az F_0 változásának irányát, az egyes megnyilatkozások hangközét, amely az alapfrekvencia-változás mértéke ($F_0\text{-max}/F_0\text{-min}$), valamint a realizáció időbeli koordinátáit. Jó viszonyítási alapnak bizonyult az egyes alapfrekvencia-értékek jellemzéséhez a beszélőnek a kijelentő mondat legvégén mért alapfrekvencia-értéke, az ún. alapvonal (vö. Varga 1998), amely egyben a szokásos kommunikációban (nem torzított hang esetén) az adott beszélő legmélyebb hangja is. Vizsgálataimban azt láttam, hogy ez az érték azonos beszélő különböző kijelentő mondataiban alig tér el egymástól. Az intenzitás mérése a kísérlet egésze szempontjából nem bizonyult relevánsnak, ami azzal magyarázható, hogy felolvasott szövegekről volt szó. A felolvasók közepes hangerővel olvastak, a hangerő változása általában csak kis mértékű volt, szignifikáns különbségek alig mutatkoztak, ezért ezt a jelen leírásban figyelmen kívül hagytam. A továbbiakban az alapfrekvencia-változás irányát értem szűkebb értelemben in-

tonáció alatt, míg a hangközt és az időtartamot kiegészítő proszódiai jegyeknek nevezem.

A felolvasott példák egyik része hümmögéseket tartalmazott, a másik pedig grammatikai szerkezettel rendelkező kérdő mondatokat. A hümmögések grammatikai szerkezettel és önálló szemantikával nem rendelkező nyelvi egységek, melyek tisztán pragmatikai funkciót hordoznak. A kérdő mondatok intonációs és proszódiai realizációja függhet a grammatikai szerkezetüktől, bizonyos szemantikai jegyeiktől (pl. beszélői attitűdök), valamint pragmatikai funkciójuktól is. Hipotézisem szerint a hümmögések és a kérdő mondatok intonációs és proszódiai realizációjának összehasonlítása értékes információkkal szolgálhat a proszódia grammatikai és pragmatikai funkcióinak elkülönítése szempontjából.

A hümmögések kontrasztív vizsgálata annál is aktuálisabb, hogy ezt a jelenséget a két nyelvben egymástól függetlenül vizsgálta Schmidt (2001) és Markó (2005). Schmidt szerint a hümmögések vizsgálatával prototipikus intonációs mintákhoz juthatunk, mégpedig vizsgálatai szerint az 1. táblázatban bemutatott 7 típushoz.

1. táblázat: Hümmögés-prototípusok és funkcióik a német nyelvben
(Schmidt 2001: 25 szerint)

| Prototípus | Funkció |
|---|---|
| 1. rövid  | tudomásulvétel |
| 2. hosszú  | időkerés, gondolkodás („Gondolkoynom kell.”) |
| 3. rövid  | befejezés jele |
| 4.  | megerősítés, egyetértés |
| 5.  | elutasítás, ellentmondás |
| 6. hosszú  | pozitív értékelés („Jó, klassz!”) |
| 7. rövid  | felszólítás válaszra, reakcióra |

Markó (2005) a magyar hümmögések öt alaptípusát különbözteti el: az 'igen' jelentésű, a 'nem' jelentésű, a figyelemjelző, a semleges kérdő és az emocionális kérdő típust. Egy magyar és nem magyar anyanyelvűekkel végzett vizsgálat pedig arra a következtetésre vezet, hogy „ezek prelingvális maradványok a kommunikációban, amelyekben sajátos monoretikus egység figyelhető meg. A nyelvfüggetlenségük pedig nyilvánvalóan motiváltságukból adódik, amely szorosan összefügg a szupraszegmentális szerkezet motiváltságának kérdésével” (103).

Mindkét kutató azt sugallja, hogy a hümmögések konvencionális intonációs mintákkal rendelkeznek, amelyek egyfajta, a pragmatika szintjén értelmezhető jelentéssel bírnak, és amelyekben az intonációs görbe és a jelentés kapcsolata nagymértékben motivált. Az én kontrasztív vizsgálataim viszont arra mutatnak, hogy a hümmögések prozódiai realizációjában a kérdő mondatokéval összehasonlítva sokkal jelentősebbek az individuális, mint a nyelvspecifikus különbségek. A kérdő mondatok esetében sokkal egységesebb, konvencionálisabb intonációs mintákat találtam, mint a hümmögéseknél. Egyes különleges pragmatikai funkcióval rendelkező kérdő mondatoknál az intonáció módosulhat, és mellé kísérő prozódiai eszközök is társulhatnak a pragmatikai töltés kifejezésére.

Markó fent idézett felfogásával alapvetően egyetértve azt szeretném kiemelni, hogy nézetem szerint a hümmögések motivált prozódiai jegyek kötegeként foghatók fel, amelyekben az egyes jegyek szabadon variálhatók, a hümmögés pragmatikai jelentését pedig alapvetően azáltal értjük meg, hogy az egyes prozódiai jegyekhez motiváltságukból adódóan az adott situációval összhangban bizonyos asszociációkat kapcsolunk. Így például egy adott kontextusban az időbeni nyújtáshoz társíthatjuk a 'gondolkodás, időkérés' funkciót, a dallam irányának többszöri megváltozásához vagy az átlagosnál sokkal nagyobb hangközhöz az érzelmi töltést, a monoton, alig változó beszéddallamhoz az érzelmileg semleges, tárgyilagos beszélői beállítottságot, az átlagosnál rövidebb realizációhoz a téma vagy a beszélgetés lezárására irányuló szándékot. Ugyanakkor az adott funkciót több prozódiai jegy is hordozhatja, és az is lehet, hogy a situációból érthetjük meg, miközben a prozódia szintjén jelöletlen marad.

Eredmények

A következőkben három hümmögéstípusnak, valamint a kérdőszavas kérdő mondatok alaptípusának és néhány pragmatikailag jelölt altípusának a német és magyar beszélők általi realizációját fogom összehasonlítani. Az eredmények az egész kísérlet szempontjából reprezentatívak.

A kiválasztott hümmögések a Schmidt által felállított tipológia 1., 2. és 6. típusát alkotják. A kísérleti személyeknek a következő dialógusrészleteket kellett felolvasni a zárójelben lévő kommentár figyelembevételével:

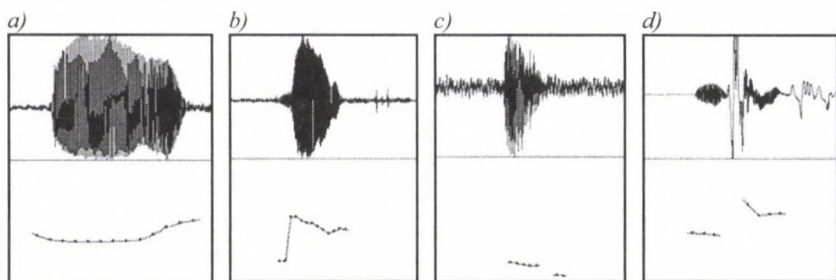
- (1) – *Erről nekem egészen más a véleményem.*
 – *Hm.* [= Értem, tudomásul vettem.]
 – *Ich bin diesbezüglich ganz anderer Meinung.*
 – *Hm.* [= Ich verstehe, ich habe es zur Kenntnis genommen.]
- (2) – *Té egyetértesz ezzel?*
 – *Hm.* [= Pillanat, gondolkodnom kell.]
 – *Bist du damit einverstanden?*
 – *Hm.* [= Moment, ich muss nachdenken.]

- (3) – *Apa, kitűnő lett a bizonyítványom!*
 – *Hm!* [Ez igen, ez már valami!]
 – *Hör mal, Vati! Ich habe die Prüfung mit sehr gut bestanden!*
 – *Hm!* [=Toll, super! Das ist schon was!]

Az első példában mindkét nyelvben igen különbözőek voltak az F_0 -realizációk. Találtam emelkedő és eső realizációkat is. A legtöbb esetben egy szótagú kiejtések voltak, amelyekben az alapfrekvencia egyenletesen ereszkedett vagy emelkedett, de néhányan két szótaggal ejtették ki a *hm* szócskát, amelyeknél az alapfrekvencia a két szótag között ugrásszerűen változott. A többi prozódiai jegy is nagy változatosságot mutat.

Közös vonás azonban mindkét nyelvű beszélők esetében, hogy a hümmögés általában egy kicsivel magasabb frekvenciatarományban realizálódik, mint a megelőző kijelentő mondat, vagyis a hümmögés legalacsonyabb pontja majdnem minden esetben valamivel magasabban van, mint a kijelentő mondat alapvonala. A másik hasonlóság a hümmögés relative kicsi hangköze volt: a hangköz minden esetben kisebb volt, mint az előző kijelentő mondat hangköze. Ez minden bizonnyal a hümmögés semleges, érzelemmentes funkciójával magyarázható. Amíg az átlagosnál nagyobb hangköz érzelmi telítettséget jelenthet, az átlagosnál kisebb az érzelem hiányára, a semleges tartalomra utalhat. A hangköz azonban német és magyar összehasonlításban különböző. Azt figyeltem meg, hogy a németek általában nagyobb hangmozgással beszélnek, az (1) példa kijelentő mondata esetében is a fókuszhangsúlyos szótag alapfrekvenciája (F_0 -max) és az alapvonal alapfrekvenciája (F_0 -min) hányadosának átlaga a németben 1,78, a magyarban 1,45 volt. Gibbon (1998) angol–német összehasonlításban megállapítja, hogy a német beszélők hangmozgása általában kisebb, mint az angoloké. Ezt a megfigyelést azzal magyarázza, hogy a szórend az angolban grammatikailag nagymértékben determinált, ezért a pragmatikai funkciókat, a mondat aktuális tagolását az intonációs eszközök jelzik. A németben a szórend részben pragmatikai funkciókat hordoz, ezért az intonációnak nem szükséges annyira egyértelműen kifejezőnek lennie. Igaz ez német–magyar vonatkozásban is, hiszen a magyar szórend grammatikailag még szabadabb, mint a német, és amint É. Kiss (1992) rámutat, teljes mértékig a mondat topik-predikátum és háttér-fókusz szerkezetét hivatott kifejezni. Mivel a magyar fókuszfrázis szórendi helye kötött, az intonáció szerepe a fókusz jelzésében redundáns. Ezért nem emeljük ki a fókusz annyira markáns intonációs eszközökkel. Ugyanakkor feltehetően ez az általában kicsi hangmozgásra való törekvés áttevődik a teljes intonációs rendszerre, így a hümmögésekre is. A hümmögés átlagos hangköze az (1) példában a németben 1,27, a magyarban 1,17 volt. Bizonyos tendenciózus különbségek mutatkoznak a hümmögés időbeni realizációjában is. A német hümmögések átlagos időtartama 0,43 s, a magyaroké 0,29 s. Viszont az egyéni realizációk esetében ez már nem kizárólagos, a legrövidebb realizáció (0,12 s) például éppen egy német beszélőtől származik. Feltehetően arról van szó, hogy a nagyobb hangmozgás

időigényesebb, ezért a német beszélőket nagyobb hangterjedelme sarkallja az általában hosszabb realizációkra, míg a magyarok többször döntenek a rövidebb kiejtés mellett.



1. ábra

Az (1) példa hümmögéseinek realizációja a németben és a magyarban. Az a) és b) német, a c) és d) magyar beszélő hangmintájából származik (Jól láthatók a nagy egyéni különbségek, míg a nyelvspecifikus különbségek csak az átlagértékek kiszámítása után tűnnek elő.)

A második példa hümmögéseiben is egyénileg igen különböző alaphfrekvencia-változásokkal találkoztam. Mind eső, mind emelkedő, mind pedig egy viszonylag magas frekvenciatartományban szinttartó görbék is előfordultak. A magyar példákban ezenkívül két eső-emelkedő intonáció is volt. A hangközérték is egyénenként igen változó. Nem meglepő ugyanakkor, hogy a beszélők túlnyomó többsége az időbeli nyújtás prozódiai eszközét használta az 'időkéres, gondolkodás' tartalom kifejezésére. A hümmögések átlagos időtartama ebben a funkcióban a németben 0,59 s, a magyarban 0,49 s volt, míg az első példában, a tudomásul vevő hümmögéseknél 0,43, illetve 0,29 s. A német hümmögések közül 11-ből 9, a magyaroknál 9-ből 8 meghaladta a 0,4 másodperces időtartamot.

A harmadik példában egy érzelmileg telített, elismerő reakciót kellett a résztvevőknek hümmögéssel kifejezniük. Az alaphfrekvencia iránya tekintetében itt is nagyok az egyéni különbségek. Mindkét nyelv esetében van eső, emelkedő, emelkedő-eső és magasan szinttartó dallam, a magyarban pedig egy alkalommal eső-emelkedő realizációval is találkoztam (2. táblázat). Gyakori az alaphfrekvencia irányának változása egy hümmögésen belül, miközben az első két példában zömében vagy egyenletesen eső, vagy egyenletesen emelkedő intonációt használtak a kísérleti személyek. A dallamirány belső változása nyilvánvalóan az érzelmi telítettséggel függ össze.

A (3) példában szereplő hümmögések másik jellemző vonása az igen nagy hangköz volt, amely szintén az érzelmi telítettség formái ismérve lehet. A 3. táblázat azt mutatja meg, hogy a német realizációk átlagos hangköze minden példánál nagyobb, mint a magyaré, ugyanakkor mindkét nyelvben jól látszik, hogy a

tudomásulvételt kifejező hümmögés átlagos hangköze meglehetősen kicsi, az érzelmdús elismerő hümmögése pedig igen nagy.

2. táblázat: Dallammenetek gyakorisága a (3) példa realizációiban

| Dallammenet | Német | Magyar |
|--------------------|-------|--------|
| eső | 3 | 2 |
| emelkedő | 3 | 1 |
| magasan szinttartó | 1 | 2 |
| emelkedő-eső | 4 | 3 |
| eső-emelkedő | 0 | 1 |

3. táblázat: A hangközök átlagértékei

| Típus | Német | Magyar |
|---------------------------------------|-------|--------|
| Fókuszhangsúlyos kijelentő mondat (1) | 1,78 | 1,45 |
| Tudomásulvételt kifejező hümmögés (1) | 1,27 | 1,17 |
| Érzelmes, elismerő hümmögés (3) | 1,84 | 1,58 |

A Student-féle t -próba segítségével ellenőrizve a tudomásulvételt kifejező és az érzelmes, elismerő hümmögés hangközértékeinek eltérése statisztikailag szignifikáns, a magyar esetben $p = 0,049$ ($df = 8$, $t = -2,312$), a német oldalon pedig $p = 0,016$ ($df = 10$, $t = -2,875$). Ugyanakkor a német–magyar összehasonlítás esetén statisztikailag nem egyértelmű az eltérés szignifikanciája, ennek bizonyítására nagyobb volumenű vizsgálat lenne szükséges.

Az időbeli realizáció tekintetében a második példában a beszélők egyértelműen élnek a nyújtás eszközével. A harmadik példa realizációi is átlagban hosszabbak, mint az első, tudomásulvételt kifejező hümmögése, ez azonban a nagyobb hangközzel és a dallammenet belső változásával magyarázható, ami nyilván több időt vesz igénybe. Német–magyar összehasonlításban a német realizációk általában hosszabbak a magyarokénál, a különbség azonban nem olyan szembe-tűnő, mint a hangközértékek esetében.

4. táblázat: Az időtartamok átlagértékei

| Hümmögés | Német | Magyar |
|-------------------------------------|--------|--------|
| (1) példa ('tudomásulvétel') | 0,43 s | 0,29 s |
| (2) példa ('időkérés, gondolkodás') | 0,59 s | 0,49 s |
| (3) példa ('érzelmes elismerés') | 0,51 s | 0,51 s |

A Student-féle t -próba tanúsága szerint a magyar oldalon az (1)–(2), illetve az (1)–(3) példa időtartam-különbsége statisztikailag rendkívül szignifikáns, a véletlen valószínűsége 1% alatt marad. A német oldalon az eltérések mértéke nem ennyire meggyőző, de az a megfigyelés, hogy a 2. példában a kísérleti személyek

gyakran élnek az időbeni nyújtás eszközével, itt is valószínűsíthető. Az (1)–(2) példák időtartamértékeit összehasonlítva a $p = 0,092$, vagyis a véletlen valószínűsége 9% körül mozog, ami nem tekinthető egyértelműen statisztikailag szignifikáns különbségnek. Az egyértelműsítéshez nagyobb mintaanagra lenne szükség.

A hűmmögések értékeléséről összefoglalóan elmondhatjuk, hogy a prelingvális motiváció egyértelműen látszik. Mindkét nyelvben nagy hasonlóságok mutatkoznak az adott funkciót kifejező választott prozódiai eszközök tekintetében, így a semleges tartalom kifejezőeszközei a viszonylag kis hangterjedelem és időtartam, az időkérsés az időbeni nyújtás, az érzelmi telítettségé a dallammenet belső változása és az átlagosnál nagyobb hangköz. Ha azonban nem átlagértékeket, hanem egyes realizációkat hasonlítunk össze, igen nagy egyéni különbségeket találunk. Ugy tűnik, hogy a hűmmögések intonációja nem vagy csak igen kevésbé konvencionizált. Azt is látjuk továbbá, hogy a hűmmögések komplex prozódiai egységek, amelyekben az intonáció, a dallammenet nem bír elsődleges szereppel, nem különíthető el a prozódia egyéb kifejezőeszközaitől. A nyelvspecifikus különbségek az individuális különbségekhez képest másodlagosak, azonban az átlagértékek összehasonlítása mutat bizonyos tendenciózus német–magyar különbségeket is, amelyek minden bizonnyal összefüggenek a két nyelv általános tipológiai sajátosságaival.

Vizsgálódásaim másik súlypontját a kérdőszavas kérdő mondatok alkották. Egyrészt azért esett a választásom erre a mondattípusra, mert az intonáció szerepe a kérdő mondatok és a kérdő beszédaktusok esetében várhatóan igen jelentős, másrészt a kérdőszavas kérdő mondatok intonációja sokkal kevésbé kutatott, mint az eldöntendő kérdő mondatoké. A szakirodalom általában csak annyit jegyez meg róluk, hogy mivel bennük a kérdőszó jelzi a kérdő funkciót, az intonáció egybeeshet a kijelentő mondatéval (eső dallammenet), de az emelkedő intonációs zárás szintén lehetséges, valamint hogy a dallamuk gyakran magasabb pontról indul, mint a kijelentő mondatoké (vö. Selting 1995: 243 kk., Olaszky 2002: 85 kk.).

Hangfelvételeim azt mutatták, hogy éppen a kérdőszavas kérdő mondatok vizsgálata szolgálhat értékes adalékkal a dolgozat elején felvetett kérdésekre. Az általam elemzett eldöntendő kérdő mondatok ugyanis egységesen a német oldalon emelkedő, a magyar oldalon emelkedő-eső intonációval zárultak, míg a kérdőszavas kérdő mondatok intonációja nagy változatosságot mutatott.

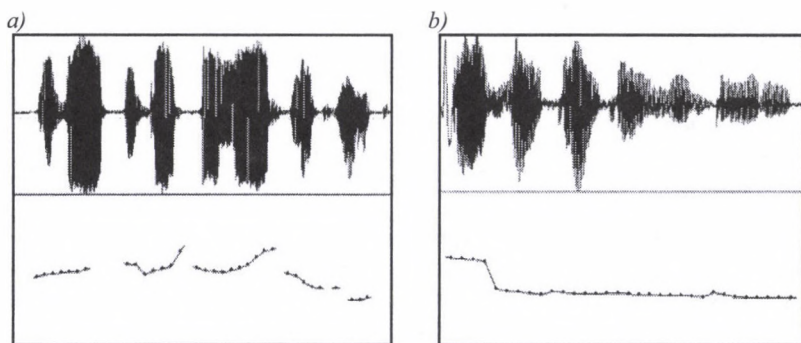
A vizsgálati eredmények jól rendszerezhetőek azáltal, ha a kérdőszavas kérdő mondatok esetében elkülönítünk egy pragmatikailag semleges alaptípust, valamint számos pragmatikailag jelölt melléktípust. Az alaptípushoz azok a tényszerű kérdések tartoznak, amelyekben a beszélő csupán egy meghatározott információt szeretne kapni a partnertől, tudásának egy hiányzó pontját szeretné kitölteni. A kérdőszó azt az információs egységet jelzi, amely a beszélő számára ismeretlen, de szeretné meg tudni. Ezekre a kérdésekre általában tehát rövid, tényszerű válaszokat vár a beszélő. Az ilyen kérdések pragmatikailag jelöletlennek tekinthetők, mert a beszélő szándéka, tehát az, hogy a partnertől információt sze-

retne kérni, levezethető a mondat szerkezetéből és szemantikai tartalmából. Ugyanakkor egy kérdési nem csak abban az esetben teszünk fel, ha egy bizonyos, általunk ismeretlen információt szeretnénk megtudni, hiszen kérdéssel kezdhethünk beszélgetést, felvethetünk új témát, vagy például ellenőrizhetjük a partner tudását (didaktikus kérdés – ebben az esetben a beszélő olyan információra kérdez rá, amely számára ismert).

A pragmatikailag jelöletlen, semleges kérdések esetében a kérdőszavas kérdő mondatok intonációja valóban megegyezik a kijelentő mondatéval, vagyis mind a németben, mind a magyarban az alapvonalra eső dallammal találkozunk. A pragmatikailag jelölt kérdések intonációja azonban ettől eltérő, nagyobb egyéni különbségeket mutat, és ezekben a beszéddallam mellett megtalálhatjuk a kiegészítő prosódiai eszközöket is.

A pragmatikailag jelöletlen, semleges kérdések esetében tehát mindkét nyelvben eső dallammal találkoztam, az alaphfrekvencia értéke a mondat elején magasabb volt, mint a mondat végén. A két nyelv között az a különbség mutatkozott, hogy a magyarban általában egyenletes volt az alaphfrekvencia esése, csökkenése, vagyis nem voltak köztes emelkedő szakaszok, míg a német mondatokban a globálisan eső alaphfrekvenciát a szó- és szakaszhangsúlyoknál kisebb emelkedő szakaszok törték meg. Ez könnyen magyarázható a két nyelv összehasonlítása kapcsán már említett alapvető különbséggel, miszerint a német intonáció fontos szereppel bír a mondat aktuális tagolásának a kifejezésében, míg a magyarban a szórendnek alapvetően pragmatikai funkciói vannak, ezért az intonáció e tekintetben kevesebb feladatot kap. Jól látható az említett különbség a (4)-es példamondat egy német és egy magyar realizációja esetében.

- (4) *Wo warst du gestern nachmittag?*
Hol voltál tegnap délután?

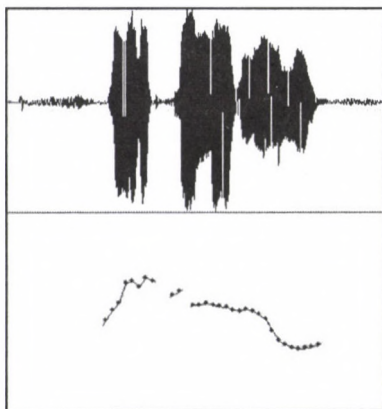


2. ábra

Egy pragmatikailag semleges kérdés (4) tipikus realizációja
 a) a németben és b) a magyarban

Didaktikus kérdések esetében a német példánál többször találkoztam azzal a jelenséggel, hogy a beszélő a mondat eleji kérdőszón belül alaphfrekvencia-emelkedéssel külön is kiemelte, hogy melyik az az egy információ, amelyikre a kérdés vonatkozik, vagyis mire vár választ a partnertől. A kérdőszó ilyen módon történő kiemeléséhez járult hozzá a nagyobb intenzitás és a magánhangzó nyújtása, amely a *wo* kérdőszó esetében szinte két szótagos realizációt eredményezett. Ehhez hasonló intonációs-prozódiai realizációt a magyar oldalon nem találtam.

- (5) *Wo liegt San Marino?*
Hol van San Marino?

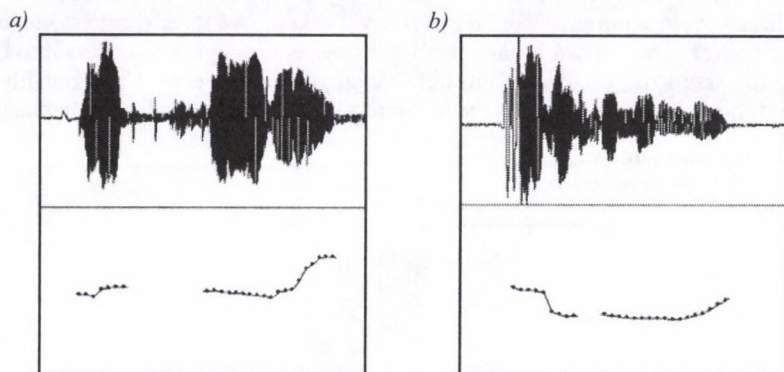


3. ábra

Didaktikus kérdés realizációja a németben (5)

Nyitott kérdésnek azokat a megnyilatkozásokat nevezem, amelyekben a beszélő kifejezi, hogy nem pusztán egy adott információ rövid, tényszerű közlését várja, hanem például egy beszélgetést kezdeményez, egy témát vet fel, a partner véleményének részletes kifejtésére kíváncsi, vitát indít, stb. A kérdőszavas kérdő mondatoknak ez egy olyan fontos melléktípusa, amelyik megfigyelésem szerint önálló, többé-kevésbé konvencionizált intonációs mintával rendelkezik. Mindkét nyelv esetében emelkedő mondatvégi dallamról van szó. A magyar realizációkban ez minden esetben úgy valósult meg, hogy az eső dallam a mondat utolsó szótagján egy kicsit visszaemelkedett, de megmaradt az alapvetően eső jellege, tehát a mondat végpontjának alaphfrekvenciája a kezdőponténál alacsonyabb értéket mutatott. A német realizációknál ugyanakkor többször előfordult, hogy az erőteljes mondatvégi emelkedés miatt globálisan emelkedő intonációba ment át az alaphfrekvencia változása, vagyis a mondat végén mért alaphfrekvencia-érték nagyobb volt a kezdőponténál. A (4)-es példának egy német és egy magyar be-

szélő által pragmatikailag nyitott kérdésként történő realizációja jól példázza ezt a különbséget.






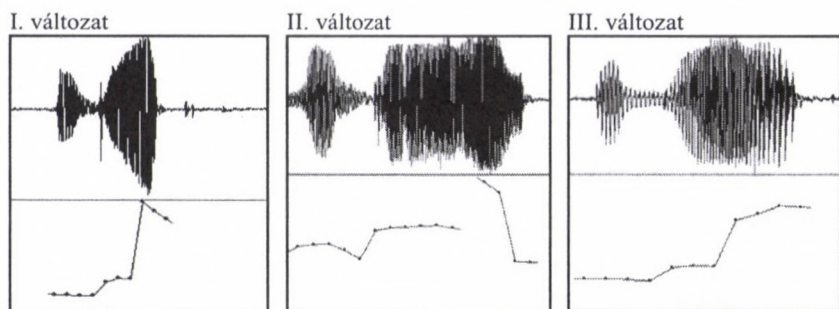
4. ábra

A (4)-es példa realizációja pragmatikailag nyitott kérdésként.
Az a) német, a b) magyar hangminta elemzése

Ugyanakkor azt figyeltem meg, hogy a magyar beszélők gyakrabban használták a nyitott kérdéseknek ezt a végén kicsit visszaemelkedő intonációs típusát, míg a német oldalon ez viszonylag ritka volt. A németek szívesebben alkalmazták a kérdőszó intonációs és prosódiai eszközökkel történő kiemelését. Ez talán azzal magyarázható, hogy mivel a németben a nagyobb hangmozgás miatt a nyitott kérdés emelkedő mondatvégi intonációja átmehet egy globálisan emelkedő intonációba, ezáltal összetéveszthetővé válik más mondattípusok intonációjával, ezért a beszélők inkább elkerülik.

Különös figyelmet érdemelnek a *Hogy vagy?* típusú kérdések, mint pl. *Mi újság? Mi történt? Mi a baj?* stb. Szinte udvariatlanság az ilyen kérdésekre rövid, lényegre törő választ adni. Az olyan válaszok, mint *Jól. Semmi. Sok minden.* azt az információt hordozzák, hogy a partner nem akar szóba elegyedni a kérdezővel. Ezek a kérdések tehát gyakorlatilag minden esetben nyitottak, pragmatikai funkciójuk nagymértékben függ az adott szituációtól. Itt hasonlóan nagy individuális különbségekkel találkoztam, mint a hümmögések esetében. Minden vizsgált esetben eltért az intonáció a semleges alaptípus eső intonációjától. A német beszélők legtöbbször egyfajta emelkedő-eső dallamot használtak, mégpedig két variációban: egy alacsony frekvenciatartományból induló, meredeken emelkedő, majd a végén kicsit visszaeső, vagy pedig az alapvonalról egy közepes frekvenciatartományból induló, az elején kicsit emelkedő, majd az alapvonalra eső dallamot. Mások egyszerű emelkedő intonációval ejtették ki a kérdést:

I. változat:  II. változat:  III. változat: 

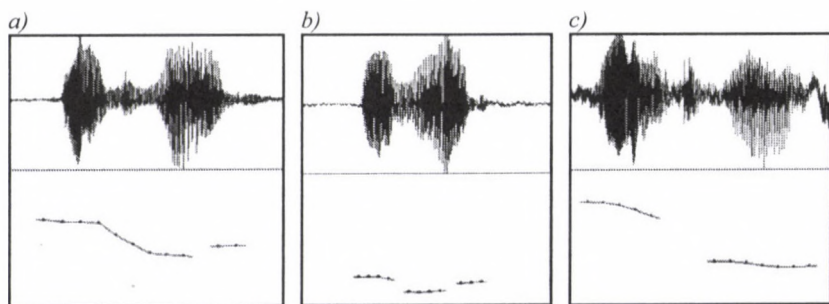


5. ábra

A *Wie gehts?* (Hogy vagy?) kérdés realizációi német hangmintákban

Jól látható, hogy a középső beszélő, aki globálisan eső dallammintát használ, a második szó magánhangzóját erőteljesen megnyújtja, vagyis a semleges intonáció mellett a pragmatikai jelöltséget egy kiegészítő prozódiai jeggyel fejezi ki.

Mivel a magyarban az emelkedő-eső dallam az eldöntendő kérdő mondatok konvencionalizált intonációja, ezekben a kérdésekben a magyar beszélők ezt nem használták. A legtöbb beszélő itt is a pragmatikailag nyitott kérdések globálisan eső, de a végén enyhén visszaemelkedő dallammintáját realizálta. Ugyanakkor nagy egyéni különbségek mutatkoztak a hangterjedelemben. Egyesek az alapvonal közelében, mások pedig magas frekvenciatarományban mozogtak. Megint másoknál az elején eső intonáció egy közepes frekvenciatarományban színtartó dallamba ment át. Egy hangmintában az intonáció az alapvonalig ereszkedett, ugyanakkor a beszélő a kérdőszó erőteljes nyújtásával és intenzitásával jelezte a kérdés pragmatikai jelöltségét.



6. ábra

A *Hogy vagy?* kérdés realizációi magyar hangmintákban

Az intonáció és a kiegészítő prozódiai jegyek összefüggését mutatja a következő magyar példa is.

(6) *Hogy írják azt, hogy folyó? J-vel vagy ly-nal?*

A legtöbb beszélő az első mondat esetében a globálisan eső, de a végén kicsit visszaemelkedő intonációs mintát használta, jelezve, hogy még egy kérdő mondat következik. Két kísérleti személy azonban egyenletesen eső dallamot realizált, ugyanakkor az utolsó szótagot erőteljesen megnyújtotta, és nagyobb intenzitással ejtette ki, ami a kérdés didaktikus jellegét fejezte ki. Megmutatták, hogy a kérdés a *folyó* szó második szótagjának írásmódjára vonatkozik.

Összefoglalás, következtetések

A hümmögések grammatikai szerkezettel és önálló szemantikai tartalommal nem rendelkező nyelvi egységek. Kizárólag pragmatikai funkciót hordoznak, melyet a prozódiai eszközök fejeznek ki. Ezért megfigyelésükből értékes nyelvészeti következtetéseket vonhatunk le az egyes prozódiai eszközök funkciójára vonatkozóan. Ugyanakkor intonációs mintáikat az én vizsgálataim szerint nem tekinthetjük az adott nyelv intonációs rendszere építőköveinek, amint azt Schmidt hipotetikusán megfogalmazta. Sokkal inkább komplex prozódiai egységekről van szó, amelyekben a motivált prozódiai eszközök a hallgatóban asszociációkat ébresztenek, ezáltal fogja a hallgató az adott szituációt is figyelembe véve felismerni a hümmögés aktuális funkcióját. Intonációs realizációjuk nagy individuális különbségeket mutat, kevésbé konvencionizált, és a nyelvek összehasonlítása esetén is csak tendenciózus különbségekről beszélhetünk, amelyek ugyan összhangban állnak a megfigyelt nyelvek tipológiai sajátosságaival, de csak több beszélőnél mért átlagértékeknél figyelhetők meg, nem kötelező erejűek. Ezzel szemben az általam vizsgált grammatikai szerkezettel rendelkező mondatok intonációja nagymértékben konvencionizált, bár a prototipikus mintáktól való eltérés segítségével itt is pragmatikai többlettartalmat lehet kifejezni. A grammatikai szerkezetek esetében az intonáción kívüli prozódiai jegyek, mint a vártnál jelentősen kisebb vagy nagyobb hangköz, a nyújtás, a rövidítés vagy az intenzitásnak a megszokottól való jelentős eltérése csak másodlagosak, és csak abban az esetben jelennek meg, ha a beszélő valamilyen sajátos pragmatikai többlettel akarja megnyilatkozását árnyalni. Összességében elmondhatjuk tehát, hogy az intonációs minták általában konvencionizáltak, és valamely grammatikai mondattípushoz rendelhetők, míg a prozódia többi eszköze kevésbé konvencionizált, motivált, individuálisan kombinálható bizonyos pragmatikai funkciók kifejezésére. Természetesen a konvencionizált intonációs minták és a motivált prozódiai jegyek nem szigorú dichotómiát alkotnak, hanem egy kontinuumot, hiszen beszéltünk erősebben és kevésbé konvencionizált mintákról, valamint, amint azt láttuk, egyes esetekben intonációs eszközök is fejeznek ki pragmatikai jelöltséget is. Feltételezhető, hogy az alapvető intonációs típusok az egyes nyelvekben a minden ember számára közös prozódiai jelrendszerből nőttek ki, miközben konvencionizálódtak, és nyelvspecifikus, grammatikai funkciók rendelődtek hozzájuk. Megkockázatom, hogy esetleg beszélhetünk egyfaj-

ta grammatikalizációs folyamatról. Ennek bizonyítása azonban további vizsgálatokat igényel.

A magyar hangmintákon egy rövid kiegészítő vizsgálatot is végeztem. Két szemináriumi csoportban meghallgattattam az összes elkészített hangmintát, és a résztvevőket megkértem, hogy tetszési indexeket rendeljenek hozzájuk, mégpedig az iskolában szokásos ötfokú osztályzási rendszerben értékeljék az egyes példákat aszerint, hogy mennyire tartják őket természetesnek vagy mesterkeltnek. Az 5. táblázat azt mutatja meg, hogy a hümmögésekre, illetve a kérdő mondatokra hány ötös, négyes stb. érdemjegy született. Az értékekből látszik, hogy a grammatikai szerkezettel rendelkező mondatok realizációit általában jobbra értékelték, mint a hümmögéseket, valamint a hümmögések értékelése szórtabb. Ezt azzal magyaráznám, hogy a hümmögések értékelésénél a kísérlet résztvevői nem tudtak konvencionálizált mintákhoz hasonlítani, így az, hogy kinek melyik hangminta tetszett vagy nem tetszett, többé-kevésbé véletlenszerű volt. Ugyanakkor a grammatikai szerkezettel rendelkező mondatok értékelésénél az értékelők feltehetően azt vizsgálták, hogy mennyire hasonlít a realizáció valamely konvencionális prototípushoz, így az értékelések egységesebbek lettek. Mivel viszonylag sokan, összesen 38-an értékelték a hangmintákat, sok hangmintát osztályoztak le, ezért a különbség szignifikáns. A statisztikában használt χ^2 -próba szerint $p < 0,01$, vagyis a véletlen szerepe 1%-nál kisebb.

5. táblázat: Tetszési indexek eloszlása
hümmögések és grammatikai szerkezettel rendelkező mondatok
intonációs-prozódiai realizációjának értékelésénél

| Osztályzat | A hümmögésekre adott osztályzatok | | A kérdő mondatokra adott osztályzatok | |
|--------------|-----------------------------------|----------------------|---------------------------------------|----------------------|
| | száma | százalékos eloszlása | száma | százalékos eloszlása |
| ötös | 172 | 39,9% | 392 | 54,0% |
| négyes | 149 | 34,6% | 207 | 28,6% |
| hármás | 75 | 17,4% | 103 | 14,2% |
| kettes | 33 | 7,6% | 16 | 2,2% |
| egyes | 2 | 0,5% | 7 | 1,0% |
| ÁTLAG | 4,05 | | 4,75 | |

A fent leírt vizsgálat, bár igen idő- és munkaigényes volt, viszonylag kis terjedelme miatt csak korlátozott következtetések levonására alkalmas. Ugyanakkor több esetben talákoztam statisztikailag szignifikáns különbségekkel, amelyek alapján elmondhatom, hogy a vizsgálatból választ kaptam azokra az alapkérdésekre, amelyek az intonáció grammatikai, illetve pragmatikai szerepével kapcsolatban a mondattípusok és mondatfajták kontrasztív összehasonlításához szükségesek. Kiderült, hogy az intonáció egyes konvencionálizált mintáihoz grammatikai funkciók rendelhetők, míg a motivált, többé-kevésbé individuálisan

használt prozódiai eszközök inkább pragmatikai funkcióval bírnak, bár közöttük átmenetek vannak.

A vizsgálatokat természetesen szisztematikusan tovább kell folytatni, hiszen sürgető a német és a magyar nyelv teljes intonációs és prozódiai rendszerének átfogó összehasonlítása. Szükség lenne autentikus beszélt nyelvi szövegekből származó példák vizsgálatára is, mégpedig különböző szituációkban. A további komplex vizsgálatok azonban csak egy nagyszabású tudományos projekt keretében képzelhetők el.

Irodalom

- Altmann, Hans 1987. Zur Problematik der Konstitution von Satzmodi als Formtypen. In Meibauer, Jörg (Hrsg.): *Satzmodus zwischen Grammatik und Pragmatik*. Niemeyer, Tübingen, 22–56.
- Altmann, Hans 1993. Satzmodus. In Jacobs, Joachim u. a. (Hrsg.): *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*. 1. Halbbd. Walter de Gruyter, Berlin–New York, 1006–1029.
- Brandt, Margareta et al. 1992. Satztyp, Satzmodus und Illokution. In Rosengren, Inger (Hrsg.): *Satz und Illokution*. Bd. 1. Niemeyer, Tübingen, 1–90.
- Cruttenden, Alan 1977. *Intonation*. Cambridge University Press, Cambridge.
- É. Kiss Katalin 1992. Az egyszerű mondat szerkezete. In Kiefer Ferenc (szerk.): *Strukturális magyar nyelvtan I. Mondattan*. Akadémiai Kiadó, Budapest, 79–178.
- Elekfi László 1962. *Vizsgálatok a hangjelentés megfigyelésének módjaihoz*. Akadémiai Kiadó, Budapest.
- von Essen, Otto 1964. *Grundzüge der hochdeutschen Satzintonation*. Henn, Ratingen.
- Féry, Caroline 1993. *German intonational patterns*. Niemeyer, Tübingen.
- Fónagy Iván – Magdics Klára 1967. *A magyar beszéd dallama*. Akadémiai Kiadó, Budapest.
- Fónagy Iván 1998. Intonation in Hungarian. In Hirst, Daniel – Di Cristo, Albert (eds.): *Intonation systems. A survey of twenty languages*. Cambridge University Press, Cambridge, 328–344.
- Gibson, Dafydd 1998. Intonation in German. In Hirst, Daniel – Di Cristo, Albert (eds.): *Intonation systems. A survey of twenty languages*. Cambridge University Press, Cambridge, 78–95.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária – Terken, Jacques 1994. Question marking in Hungarian: Timing and height of pitch peaks. *Journal of Phonetics* 22. 269–281.
- Hetland, Jorunn – Molnár Valéria 1995. Zur Funktion der Prosodie in Fragen. Eine kritische Diskussion mit Ausgangspunkt in Margret Seltings Aufsatz „Phonologie der Intonation“ (1993). In Önnerfors, Olaf (Hrsg.): *Festvorträge anlässlich des 60. Geburtstags von Inger Rosengren*. (Sprache & Pragmatik, Sonderheft). Lund, 3–46.
- Isačenko, Alexander V. – Schädlich, Hans-Joachim 1971². Untersuchungen über die deutsche Satzintonation. In Bierwisch, Manfred (Hrsg.): *Untersuchungen über Akzent und Intonation im Deutschen*. Studia Grammatica VII. Berlin, 7–67.
- Keszler Borbála (szerk.) 2000. *Magyar grammatika*. Tankönyvkiadó, Budapest.

- Kohler, Klaus J. 1991. A model of German intonation. In Kohler, Klaus J. (ed.): *Studies in German intonation*. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel 25. Kiel, 11–27.
- Markó Alexandra 2005. „Szavak nélkül.” Nonverbális vokális közlések fonetikai elemzése. *Magyar Nyelvőr* 129/1. 88–104.
- Olaszy Gábor 2001. Prozodémák fonetikai reprezentációja. *Beszédkutatás* 2001. 28–45.
- Olaszy Gábor 2002. A magyar kérdés dallamformáinak és intenzitás szerkezeteinek fonetikai vizsgálata. *Beszédkutatás* 2002. 83–99.
- Pierrehumbert, Janet B. 1980. *The phonology and phonetics of English intonation*. Garland, New York.
- Pheby, John 1980. *Intonation und Grammatik im Deutschen*. Zweite Auflage. Akademie Verlag, Berlin.
- Rabanus, Stefan 2001. *Intonatorische Verfahren im Deutschen und Italienischen. Gesprächsanalytische und autosegmentale Phonologie*. Niemeyer, Tübingen.
- Schmidt, Jürgen Erich 2001. Bausteine der Intonation? In Schmidt, Jürgen Erich (Hrsg.): *Neue Wege der Intonationsforschung*. Germanistische Linguistik 157–158. Georg Olms Verlag, Hildesheim–Zürich–New York, 9–32.
- Selting, Margret 1995. *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Niemeyer, Tübingen.
- Varga László 1998. Dallamelemek és szótagok asszociációja a magyar hangjelzés autosegmentális tárgyalásában. *Általános Nyelvészeti Tanulmányok XIX.* 257–272.
- Varga László 2001. Az eső-emelkedő hangjelzésforma a magyar nyelvben. *Beszédkutatás* 2001. 17–27.
- Varga László 2003. Dallami meghatározottságok a magyar mondatban (a topik dallama). *Beszédkutatás* 2003. 147–162.
- Wunderlich, Dieter 1988. Der Ton macht die Melodie. In Altmann, Hans (Hrsg.): *Intonationsforschungen*. Linguistische Arbeiten 200. Niemeyer, Tübingen, 1–40.

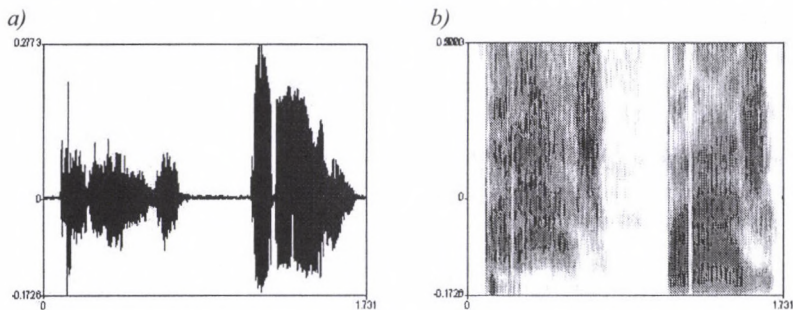
SUTTOGOTT EJTÉSŰ MAGÁNHANGZÓK ÉSZLELÉSE

Gráci Tekla Etelka

Bevezetés

A suttogás a beszédképzés természetes módja, bármikor alkalmazhatjuk, és bármikor hallhatjuk is, amennyiben a beszédhelyzet megkívánja a beszédpartner megítélése szerint. Kommunikációnk során tehát szükségünk lehet suttogott ejtésű beszéd feldolgozására, azonban ezen beszédmod az artikuláció tekintetében a hangszalagok működésében, az akusztikai szerkezet tekintetében pedig több paraméterben is eltér a normál fonációjú beszédjeltől, ami a beszédpercepcióra is hatással lehet.

Ezen akusztikai eltérésekkel több tanulmány is foglalkozott, eredményeiket Morris–Clements (2002) foglalta össze. A következő főbb vonások jellemzik a suttogott beszédet a normál ejtésűvel szemben (1. ábra). A hangforrás a vokális traktus alsó részében van, így a beszédjel jelentősen (mintegy 20 dB-lel) alacsonyabb nyomású, mint a normál fonációjú beszéd, 500 Hz alatt a spektrum elhalványodik, 500 és 2000 Hz között pedig tipikusan kevésbé változatos. A beszédhangokra jellemző akusztikai szerkezet megmarad, azonban eltérhet a normál fonációjú beszédben tapasztalható értékektől: a magánhangzók esetében például az első formáns elhalványodik, illetve a formánsértékek szignifikánsan (főként az elsőé) megemelkednek (vö. még Gósy 2002). Mindezen túl a hangsor intenzitásviszonyai is megváltoznak, míg a mássalhangzók intenzitását kevésbé befolyásolja a fonáció módja, a magánhangzókét erőteljesebben csökkenti.



1. ábra

A *parázs* szó a) oscillogramja és b) spektrogramja ugyanazon beszélő suttogott (bal oldalon) és normál (jobb oldalon) ejtésében

Hétköznapi tapasztalataink alapján elmondhatjuk, hogy a suttogott ejtésű beszéd észlelése nehezebbnek hathat a normál fonációjával szemben. Kérdés lehet azonban, hogy ezt a nehézséget a gyakoribb, így megszokottabb beszédmódhoz képest módosult akusztikai kulcsok okozzák-e. A mássalhangzók zöngésségi azonosítása esetében azt tapasztaltuk, hogy ugyan az észlelés bizonytalan, az azonosítás mégis jobb eredményeket mutat a véletlen találathoz (Grácz 2005). Felmerülhet a kérdés, hogy hogyan működik a percepció a magánhangzók esetében, amikor az elsődleges megkülönböztető jegy nem neutralizálódik, csak módosul.

A magánhangzók észlelése

A magánhangzók legelfogadottabb észlelési modellje szerint a formánsok egymáshoz viszonyított aránya alapján azonosítjuk a magánhangzókat (Strange 1999). A mai napig több modell kidolgozásával foglalkoznak – felmerült például az F_0 szerepe az azonosításban (MAS Model of Vowel Perception – Stokes é. n.), illetve az a lehetőség, hogy a feldolgozás során az elhangzó hangsort rövid, néhány milliszekundumnyi egységekre osztva hasonlítjuk össze a lehetséges neurális spektrogramokkal (Hillebrand–Houde 2002).

A formánsok arányán alapuló feldolgozási modell válasz lehet a különböző személyek ejtése közötti különbség feldolgozására, hiszen a hangképző szervek változatossága miatt a formánsok nagyon eltérőek lehetnek beszélőnként (vö. pl. Gósy Mária 1984-es gyermeknyelvi vizsgálatával), arányuk azonban nagyobb hasonlóságot mutat a személyek között.

Mivel a suttogott ejtésű beszéd esetében a magánhangzók formánsértékei szignifikánsan megemelkednek, azonban az első formáns nagyobb mértékben (Gósy 2002), lehetséges, hogy ez a módosulás olyan fokú, hogy a magánhangzók azonosítása sérül. Assmann–Nearey (é. n.) mesterségesen módosított hanganyagban végzett kísérleteinek eredménye azt mutatta, hogy a formánsértékek változása megnehezíti a hang azonosítását.

A beszédhangok azonosításával kapcsolatban további kérdéseket vethet fel a koartikuláció okozta szerkezeti módosulás. A jelen kísérletben a nazális mássalhangzó-környezet hatását vizsgáltuk, mivel a magánhangzók nazalizálódásakor a lágy szájpad leereszkedik, ami az egyes magánhangzók képzésekor különböző mértékű hatást eredményez az első, második és a harmadik formánsban – eltérő mértékben (Ladefoged 2005). Horváth Viktória (2005) kimutatta, hogy a nazalizáció hatásának mértéke nemcsak az egyes beszédhangok, hanem a beszélő és az adott nazális függvénye is, illetve hogy az egyes formánsok a magyarában is eltérő mértékben változnak (emelkednek meg). Ez ismét azt jelenti, hogy a formánsok arányai jelentős mértékben módosulhatnak. Kísérletünkben tehát azt is vizsgáltuk, hogy a suttogott ejtésű, nazális kontextusban szereplő vokálisok azonosítása a feltételezett nagyobb mértékű akusztikai változás következtében hogyan alakul.

Kísérletünk hipotézisei a következők:

a) A suttogott ejtésű magánhangzók felismerése a megszokottól eltérő akusztikai szerkezet miatt nehézségekbe ütközik, bizonytalan.

b) Feltételezzük, hogy a nyelv függőleges helyzetének téves észlelése lesz a leggyakoribb – egyrészt mivel normál ejtésű beszéd esetében a téves azonosítások 98%-a a nyelv függőleges helyzetének hibás észlelésének következménye (Stokes é. n.), másrészt pedig Gósy Mária (2002) adatai alapján, miszerint suttogott ejtésben az első formáns szignifikánsan magasabbra emelkedik, mint a második. Ezen belül pedig a felsőbb nyelvállásfokú hangok alsóbb nyelvállásfokúként való azonosítását tételezzük jellemzőbbnek, mivel az első formáns értéke az artikulációs tényezők közül a nyelvállásfokkal van összefüggésben úgy, hogy az alacsonyabb nyelvállásfok magasabb F_1 -értékhez vezet az akusztikumban (Gósy 2004).

c) Feltételezésünk szerint a nazális környezetben szereplő magánhangzók azonosítása további nehézségeket okoz az adatközlők számára.

Anyag és módszer, kísérleti személyek

A hipotézisek igazolására egy kísérletsorozatot állítottunk össze, amelyhez a hanganyagot egy női adatközlővel, csendesített szobában, minidisc-felvevővel rögzítettük. A percepció tesztet 39 egyetemistával: 20 nővel és 19 férfival végeztettük el. Átlagéletkoruk 21,6 év (1. táblázat).

1. táblázat: Az adatközlők életkori jellemzői

| | Átlagéletkor (év) | Szóródás (év) |
|---------|-------------------|---------------|
| Nők | 21,4 | 19–28 |
| Férfiak | 21,8 | 18–25 |

A teszt két altesztből, logatomok (jelentés nélküli hangsorok) és szavak magánhangzó-minőségeinek megítéléséből állt. A kétféle tesztelés oka, hogy a logatomok esetében ugyan minden vizsgálni kívánt fonetikai helyzetet tudtunk elemezni, a szavak esetében pedig nem, viszont a logatomok észlelése nehezebb feladat a kísérleti személyek számára, mint a jelentéssel bíró egységeké, ezért a két részeszt eredményei összevetve pontosabb képet adhatnak a percepció működéséről.

Az első feladatban tehát 36 logatomot hallottak a kísérleti személyek. Ezekben kilenc magánhangzó minden fonetikai helyzetben, vagyis abszolút szó eleji, szó belseji (mind a megelőző, mind a követő hang mássalhangzó volt), illetve abszolút szóvégi helyzetben négyszer szerepelt. Ezek közül két esetben a szomszédos mássalhangzó(k) nazális(ok) volt(ak), két esetben pedig nem nazális hangok. A logatomok mindegyike így három szótagból állt, melynek felépítése VC(C)VC(C)V volt. Pl.: *üncöge, önyinzé, arifó, utnáme, égüvä*. A tesztlapon csak a hangsorok mássalhangzóit szerepelték (pl. *_nc_g_*), és arra kértük az adat-

közlőket, hogy egyszeri hallás után pótolják a hiányzó magánhangzóknek megfelelő betűket.

A második feladatban 30 szó hangzott el. A tesztlapon szópárokat, szócsoportokat adtunk meg lehetőségnek. A szópárok, szócsoportok olyan minimálpárok-ból álltak, melyek csak egy magánhangzóban tértek el egymástól. Ebben az altesztben a magánhangzók csak szó belseji, két mássalhangzó közötti helyzetben szerepeltek. A szópárok minden esetben egy-egy képzésjegyben eltérő vokálist tartalmaztak (félkövérrel jelöljük a lejátszott tesztszót), pl.: a nyelv függőleges helyzete: **víg-vég**, a nyelv vízszintes helyzete: **sor-sör**; kerekítettség: **kép-köp**, míg a szócsoportok minden esetben több képzésjegyben eltérő hangokat is tartalmaztak, pl. **hit-hat-hát-hét-hőt-hűt**, **tűr-tar-tör-tér-túr-tár**. Az adatközlők feladata az volt, hogy az egyes szópárok/szócsoportok tagjai közül kiválasszák, melyik hangzott el. (A feladatban azért adtunk meg választási lehetőséget a kísérleti személyeknek, hogy ezzel csökkenthessük az adott szavak közötti gyakorisági eltérés hatását.) Ebben a feladatban nem szerepeltek nazálisok, mivel kevés olyan minimálpár létezik, amelyben mindkét mássalhangzó nazális.

A kísérletben nem vettük figyelembe a hosszúsági oppozíciót, vagyis nem tettünk különbséget hosszú-rövid hangpárok között, mivel azok elsődleges megkülönböztető jegye az időtartam (Gósy 2004), aminek vizsgálata a jelen dolgozatban nem volt célunk.

Az adatokon a 9.0 verziós számú SPSS szoftverrel végeztünk statisztikai elemzéseket, egytényezős varianciaanalízist, valamint az egyes csoportok egymáshoz való viszonyának vizsgálatához a Tukey-féle post hoc tesztet alkalmaztuk. A próbákat minden esetben 95%-os szignifikanciaszinten végeztük.

Eredmények

A kapott válaszokat mindkét feladat esetében értékeltük a magánhangzók képzési jegyeinek tévesztésének függvényében, vagyis a nyelv függőleges és vízszintes mozgása, valamint az ajakműködés szerint. Ezen belül pedig a téves észlelés iránya alapján, azaz hogy az elhangzott hang helyett annak nyelvéllásfokától alacsonyabb vagy magasabb nyelvéllásfokú hangot, labiális magánhangzó helyett illabiálisat, vagy illabiális helyett labiálisat, valamint hogy veláris hang helyett palatális vagy palatális hang helyett veláris hang észlelése jellemzőbb. Megvizsgáltuk továbbá, hogy az elhangzó hang nyelvéllásfokának függvényében hogyan alakulnak az észlelési eredmények. A logatomok esetében pedig megnéztük a mássalhangzó-környezet nazalitásának és a vizsgált hangok fonetikai helyzetének hatását az észlelésre.

Az értékeléskor mindig az adott kategóriát tekintettük 100%-nak, vagyis pl. a nyelvéllásfok tévesztésének vizsgálata esetén azon magánhangzók, amelyeknél létezik alsóbb nyelvéllásfokú magánhangzó (vagyis az [a:] nem tartozik ide), illetve azok, amelyeknél létezik magasabb nyelvéllásfokú magánhangzó (vagyis az [i, u, y] nem tartozott ide), külön-külön jelentettek 100%-ot.

A magánhangzók észlelési eredményei a suttogott logatomokban

A logatomok esetében tehát megvizsgáltuk, hogy 1. a fonetikai helyzet hatással van-e a magánhangzók észlelésére; 2. mely képzésjegy(ek) észlelése ütközik nagyobb nehézségbe suttogott hangsorok esetében; 3. mik az egyes képzésjegyek tévesztésén belüli esetleges tendenciák; 4. melyek a leginkább tévesztett magánhangzó-kategóriák; illetve hogy 5. a nazálisok koartikulációs hatása okoz-e további nehézségeket a kísérleti személyek számára.

A kapott eredmények azt mutatják, hogy a magánhangzók észlelése kevesebb, mint egytöd részben téves (2. táblázat). Vagyis azt mondhatjuk, hogy nagymértékben jól azonosíthatóak a magánhangzók a suttogott beszédben annak ellenére is, hogy a hangsorok jelentés nélkülsége megnehezíti a percepció működését.

2. táblázat: A logatomok esetében kapott téves válaszok arányai

| Hibaátlag | Hangtévesztés (az összes hang %-ában) |
|-----------|--|
| Nők | 18,90 |
| Férfiak | 16,16 |
| Összesen | 17,56 |

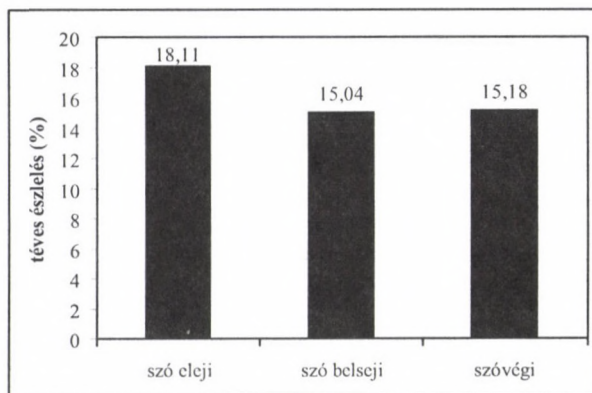
A hangsorbeli helyzet függvényében a kísérleti személyek valamivel nagyobb arányban tévesztették a hangsor eleji magánhangzókat, mint a hangsor belsejüket, illetve -végüket (2. ábra), azonban ez a különbség statisztikailag nem szignifikáns (a post hoc teszt eredményei: $F(3, 114) = 2,551$; $p = 0,117$, $p = 0,140$, $p = 0,996$). Feltételezhető, hogy a hangsor eleji hangok valamivel magasabb tévesztésének a figyelem még nem kellő mértéke lehet az oka, vagyis hogy a kísérleti személyek ekkor kezdenek el figyelni az elhangzó hangsorra.

Megvizsgáltuk, hogy a magánhangzók mely képzésjegyét/eit érinti leginkább a tévesztés (3. ábra). A logatomok azonosításakor az elhangzó hang képzésjegyeit tekintve a nyelv függőleges helyzete látszik a legnehezebbnek az észlelés számára, tévesztése meghaladja a 10%-ot, míg a nyelv vízszintes helyzete és a kerekítettség tévesztése nem éri el a 0,5%-ot. (Ez azt jelent, hogy a nyelv vízszintes helyzetét egy vagy két esetben tévesztette el összesen 18 adatközlő, a kerekítettséget pedig összesen hatan tévesztették egy-egy esetben. Ezekkel ellentétben a nyelv függőleges helyzetének téves észlelése minden adatközlőnél előfordult, gyakran több mint 10 esetben is.) A nyelv függőleges helyzete bizonyult a legnehezebben azonosítható képzésjegynek, ez a különbség szignifikáns ($F(4, 154) = 218,57$; $p < 0,001$).

Tíz esetben a kísérleti személyek nem egy nyelvállásfokbeli szintet tévesztettek (pl. [u] → [ɔ]). Ez az összes ilyen típusú tévesztési lehetőség arányában 0,27%.

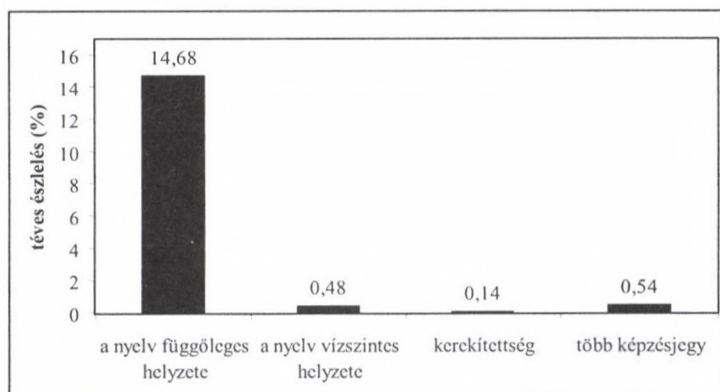
A több képzésjegybeli tévesztés kategóriáját jobban megvizsgálva azt találjuk, hogy kevesebb, mint a kísérleti személyek felénél fordult elő (17 a 39-ből), sze-

mélyenként egy, illetve két esetben (a több képzésjegybeli téves észleletek mintegy 41%-ában, míg háromszor csak egy adatközlőnél találtunk több képzésjegybeli tévesztést, ami 2,56%-ot jelent). A nyelvállásfokbeli tévesztés itt is kimagasló volt, a 23 adatból csak egyben nem találtunk ilyen eltérést az elhangzott és az észlelt hang között.



2. ábra

A logatomokban vizsgált magánhangzók téves észlelése a fonetikai helyzet függvényében

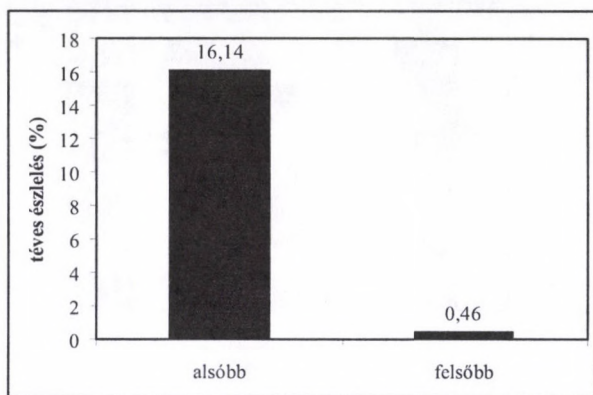


3. ábra

A logatomokban vizsgált magánhangzók téves észlelése a képzésjegy függvényében

Az egyes képzésjegybeli tévesztésekre a következő eredményeket kaptuk. A nyelv függőleges állása szerinti tévesztés esetében az elhangzott hang helyett

alsóbb nyelvválásfokú hang észlelése 30-szor gyakoribb, mint a felsőbb nyelvválásfokú hang azonosításának esete (4. ábra). Ez statisztikailag szignifikáns eltérés ($F(2, 76) = 252,61$; $p < 0,001$). A több képzésjegybeli tévesztések esetében valamivel kisebb az eltérés a két tévesztési irány között, a 22 téves észlelet közül, amelyekben nyelvválásfokbeli tévesztés is történt, 18-szor alsóbb, és csak 4-szer felsőbb nyelvválásfokú az észlelt hang az elhangzónál.

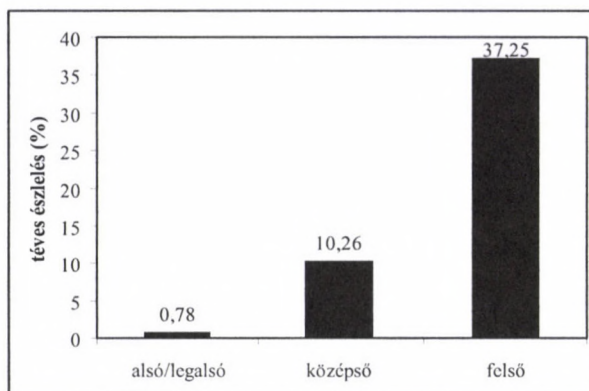


4. ábra

A logatomokban vizsgált magánhangzók tévesztésének tendenciái az elhangzott hang nyelvválásfokának függvényében

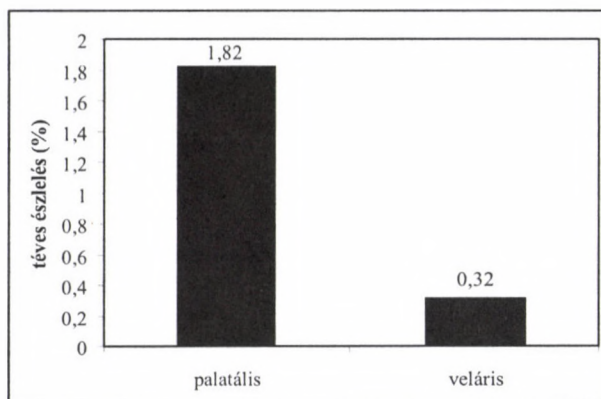
Az elhangzott hang nyelvválásfoka szerint megvizsgálván az adatokat a következő eredményeket kaptuk (5. ábra). A felső nyelvválásfokú magánhangzók észlelése okozta a legtöbb nehézséget (mintegy 37%-ban tévesztették), míg a középső nyelvválásfokúakat csak 10%-ban tévesztették, illetve az alsó/legalsó nyelvválásfokú magánhangzókat 0,8%-ban. Ezek az észleletbeli eltérések szignifikáns különbséget mutattak, $p < 0,001$ mind az ANOVA, mind a Tukey-féle post hoc tesztek esetében ($F(3, 114) = 142,39$). Érdeemes megjegyezni, hogy az [ε] hangot minden esetben helyesen azonosították, más elhangzó hang helyett azonban észleltek [ε] hangot.

A nyelv vízszintes mozgásának tévesztése esetén (6. ábra) jellemzőbb a veláris hangok téves azonosítása (a palatális hangok velárisként való észlelése 3-szor fordult elő, míg a veláris hangok palatálisként való észlelése 17-szer, vagyis majdnem hatszor gyakrabban). Ismét azt tapasztalhatjuk a több képzésjegybeli tévesztésekkel összehasonlítva, hogy az utóbbiaknál kisebb a két tévesztési irány aránya, azonban ebben az esetben is hasonló marad a tendencia, azaz a 18 esetből, amikor a palatalitás észlelése hibás, 13-szor (72,22%) veláris hangot észlelnek palatálisként, és csak 5-ször (27,78%) palatális hangot velárisként – tehát közel háromszor annyi esetben azonosították tévesen a veláris magánhangzókat.



5. ábra

A logatomokban szereplő vizsgált magánhangzók téves észlelése az elhangzott hang nyelvállásfokának függvényében

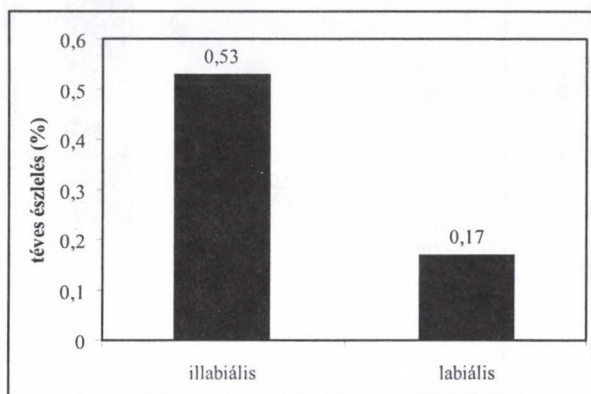


6. ábra

A logatomokban vizsgált magánhangzók tévesztésének tendenciái a nyelv vízszintes mozgásának függvényében

A kerekítettségbeli tévesztések esetében ismét eltérést láthatunk a két lehetséges tévesztési irány között (7. ábra). Illabiális magánhangzót labiálisnak összesen egyszer, labiálisat illabiálisnak ötször észleltek a kísérleti személyek. Ez a különbség azonban statisztikailag nem szignifikáns ($F(2, 76) = 2,923$; $p = 0,091$). A több képzségjegybeli tévesztések esetében ellentétes irányú tendenciát láthatunk, vagyis a hat esetből, amikor az elhangzott hang kerekítettségét is tévesen észlelik

az adatközlők, négy esetben azonosítanak illabiális hangot labiálisnak, míg ellenkező irányú téves észlelés kétszer fordul elő.



7. ábra

A logatomokban vizsgált magánhangzók
kerekítettségbeli tévesztésének tendenciái

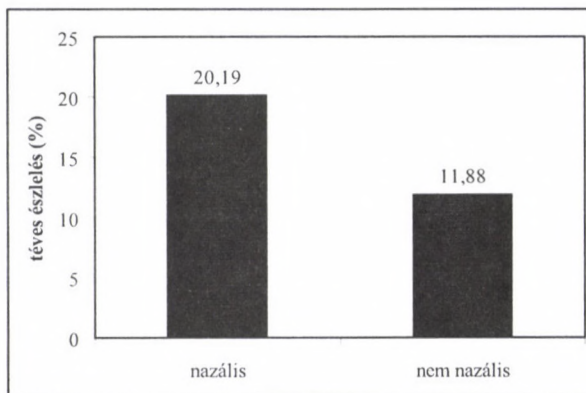
A logatomok esetében megvizsgáltuk a nazalizáció hatását az észlelésre, okoz-e további nehézséget az adatközlők számára, és ha igen, milyen mértékben (8. ábra). Az eredmények a következőképpen alakultak. A nazális környezet hatására mintegy 20%-ban észlelték tévesen az adatközlők az elhangzott hangot, míg egyéb mássalhangzók környezetében mintegy 12%-ban, vagyis az eltérés közel kétszeres, szignifikáns különbség ($F(2, 76) = 29,76; p < 0,001$). Megállapíthatjuk tehát, hogy a nazalizáció megnehezíti a suttogott ejtésű magánhangzók észlelését.

Összegezve a logatomokban elhangzott magánhangzókra kapott eredményeket: a magánhangzók azonosítása nem okozott jelentős nehézséget a kísérleti személyek számára. Az észlelés számára a magánhangzók képzéssjegyei közül egyértelműen a nyelv függőleges helyzete bizonyult a legnehezebben azonosíthatónak, ezen belül pedig az elhangzott hang alsóbb nyelvvállásfokúként való azonosítása mutatkozott a leggyakoribb észlelési hibának. Az egyéb képzéssjegyek téves észlelésében is kimutathatóak egyértelmű téves azonosítási tendenciák.

Az [ɛ] hangot egy esetben sem észlelték tévesen az adatközlők, ami valószínűleg összefügghet az adott hang gyakoriságával. Gósy Mária (2004) kutatásai szerint a spontán beszédben az összes előforduló magánhangzó arányában 28%-ban fordul elő.

Adataink alapján a felső nyelvvállásfokú magánhangzókat volt a legnehezebben azonosítani az adatközlők számára. A legkevesebbszer az [y] hangot azonosították helyesen a kísérleti személyek, az esetek 57,05%-ában tévesen észlelték, ez

pedig a legritkábban előforduló magánhangzó (ritkább, mint 2%) Gósy Mária (2004) adatai alapján.



8. ábra

A logatomokban vizsgált magánhangzók téves észlelése a mássalhangzó-környezet nazalitásának függvényében

A magánhangzók észlelési eredményei a suttogott szavakban

A szavak esetében ugyancsak megvizsgáltuk, hogy 1. a magánhangzók mely képzésjegye(i) a legnehezebben azonosítható(ak); 2. a tévesztések „irányát”; illetve 3. hogy az elhangzó magánhangzók függőleges nyelvállásfoka alapján mely hangcsoport a legnehezebben észlelhető suttogott ejtésben.

A kapott eredmények azt mutatják, hogy a szavak esetében a magánhangzók téves észlelése nem jelentős, az összes hang 2%-ánál kevesebb esetben azonosították hibásan az adatközlők a vizsgált hangot (3. táblázat).

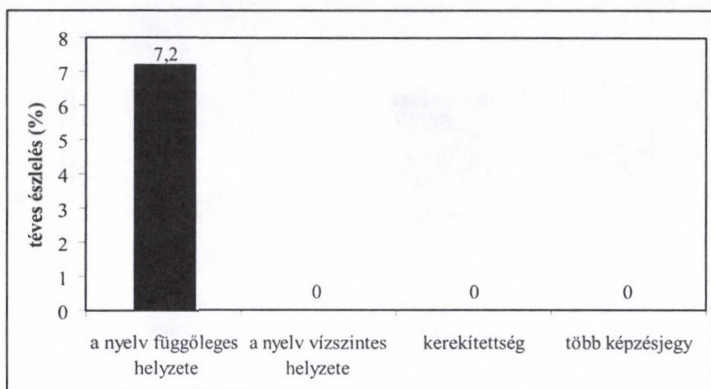
3. táblázat: A szavak esetében kapott téves válaszok arányai

| Hibaátlag | Hangtévesztés (az összes hang %-ában) |
|-----------|--|
| Nők | 1,65 |
| Férfiak | 1,42 |
| Összesen | 1,54 |

Az eredményeket a magánhangzók képzésjegyei alapján csoportosítva láthatjuk, hogy a kísérleti személyek csak a nyelv függőleges helyzetét tévesztették a szavakban vizsgált magánhangzók esetében (9. ábra).

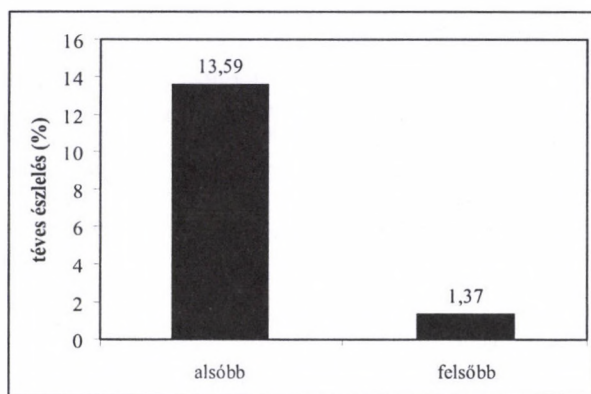
A nyelv függőleges állásának tévesztése esetén ismét az elhangzó magánhangzóhoz képest alacsonyabb nyelvállásfokú hangot észleltek a kísérleti sze-

mélyek, az ellenkező irányú tendenciánál mintegy 90%-kal gyakrabban (10. ábra), és ez a különbség szignifikánsnak bizonyult ($F(2, 76) = 77,17; p < 0,001$).



9. ábra

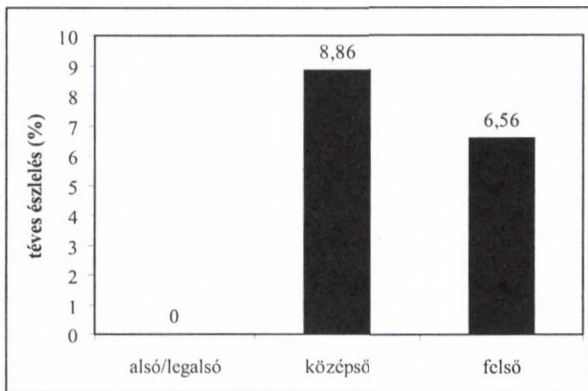
A szavakban vizsgált magánhangzók téves észlelése a magánhangzók képzőjeinek függvényében



10. ábra

A szavakban vizsgált magánhangzók téves észlelésének tendenciái a nyelv függőleges állása alapján

Az alsó/legalsó nyelvállásfokú magánhangzókat egy esetben sem észlelték tévesen az adatközlők, míg a középső és a felső nyelvállásfokú magánhangzók nagyjából azonos nehézséget okoztak számukra (11. ábra). A három csoport közötti különbség szignifikáns, mindhárom esetben $p \leq 0,001$ ($F(3; 114) = 47,78$).



11. ábra

A szavakban vizsgált magánhangzók téves észlelése az elhangzott hang nyelválásfokának függvényében

Az eredmények összegzéseként elmondhatjuk, hogy a vizsgált szavak esetében a nyelválásfokot tévesztették csak az adatközlők, és tendenciájában ismét alacsonyabb nyelválásfokú magánhangzóval azonosították az elhangzott hangot.

Ebben a feladatban az [o], [e:], [i] hangok azonosítása okozott nehézséget. Egyéb hang tévesztése ([ø]) egyszer fordult elő. (Ez a következő tévesztéseket jelenti: *szét* → *szít*, *víg* → *vég*, *rojt* → *rajt*, *főz* → *füz*.)

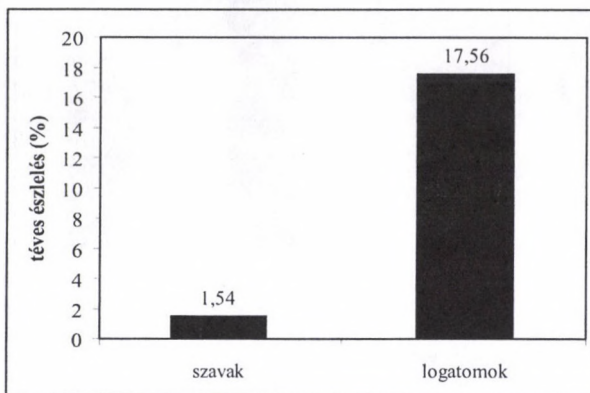
A magánhangzók észlelési eredményeinek összegzése

A két feladat között jelentős különbséget láthatunk a magánhangzó-észlelés bizonyosságában, mintegy tízszer több hibás választ adtak a logatomokban szereplő magánhangzók azonosításakor a kísérleti személyek (12. ábra). Mindazonáltal egyik feladatban sem mondható nagymértékben bizonytalannak a magánhangzó-azonosítás. A különbséget valószínűleg a két teszt módszere közötti eltérés okozta, vagyis az, hogy a logatomok azonosítása nehezebb feladat a percepció számára, mint a jelentéses egységeké, valamint a logatomok esetében nem adtunk meg választási lehetőséget a kísérleti személyeknek.

A téves észleleteket a képzségjek mentén összevetve a következőt találtuk (13. ábra). Egyértelműen mindkét feladatban szignifikánsan a nyelv függőleges helyzete bizonyult (a leggyakrabban) hibásan azonosított képzségjegynek. A szavak esetében nem is jelenik meg egyéb típusú tévesztés, tehát valószínű, hogy a nyelv vízszintes állásának és a kerekítettségnek a tévesztése valóban ritka, illetve a logatomazonosítási feladat nehézsége idézhette elő.

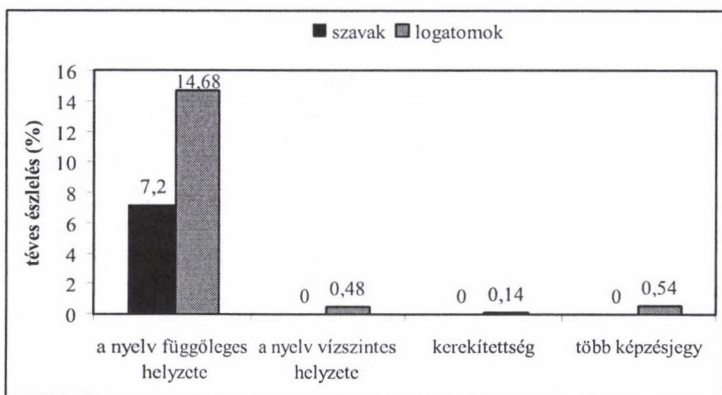
A nyelv függőleges helyzetének tévesztési „iránya” mindkét feladattípus függvényében azt mutatta, hogy az észlelt hangnál alacsonyabb nyelválásfokú hangot azonosítanak hibás észlelés esetén nagyobb valószínűséggel az adatköz-

lők, mint felsőbbet. Ugyan az eltérés valamivel kisebb a szavak esetében, de még itt is közel tízszeres (14. ábra).



12. ábra

A vizsgált magánhangzók téves észlelése a feladat típusának függvényében

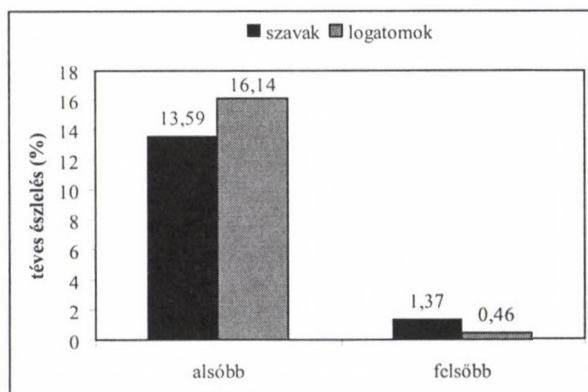


13. ábra

A vizsgált magánhangzók téves észlelése a képzésjegy és a feladat típusának függvényében

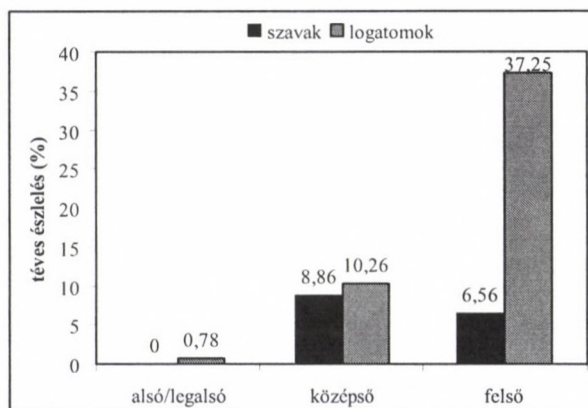
Az elhangzott hang alapján megvizsgálva az adatokat jelentős eltérést találtunk a két feladattípus között (15. ábra). A szavak esetében mintegy egyharmadnyival többször azonosították tévesen a középső nyelvválásfokú magánhangzókat, mint a felső nyelvválásfokúakat, ugyanakkor a logatomok esetében ez az eltérés fordított és nagyobb arányú, vagyis a felső nyelvválásfokú magánhangzók észlelése bizonyult közel ötször nehezebbnek. Ezt az eltérést a két feladat ered-

ménye között az egyik kísérleti szó realizációja okozhatta. A *rojt* szó /j/ fonémája suttogott ejtésű realizációjának koartikulációs hatása lehetett olyan mértékű, ami miatt az /o/ fonéma inkább az [ɔ] hanghoz közeli ejtésben valósult meg.



14. ábra

A vizsgált magánhangzók téves észlelésének tendenciái a nyelvváltsfok és a feladat típusa függvényében



15. ábra

A vizsgált magánhangzók téves észlelése az elhangzott hang nyelvváltsfokának és a feladat típusának függvényében

Összegezve: a logatomészlelési feladatban a legnehezebb a felső nyelvváltsfokú magánhangzók azonosítása, azonban jelentéss nyelvi egységek esetén ezen hangzócsoportot is kevesebb esetben tévesztik a kísérleti személyek.

Következtetések

A kísérlet eredményei alapján a suttogott ejtésű magánhangzók azonosítása esetén a nyelvvállásfok tévesztése volt a leggyakoribb hiba, ezen belül pedig a felsőbb nyelvvállásfokú magánhangzók helyett alsóbb nyelvvállásfokú észlelése jellemző. Vagyis összefüggést találhatunk az első formáns nagyobb mértékű emelkedése és a hibás észleletek között.

A mássalhangzó-környezet nazalitása további hatással volt a magánhangzó-azonosításra, több téves választ kaptunk a nazális kontextusban szereplő, mint az egyéb konzonánsok között elhangzó vokálisokra.

Érdemes megjegyezni, hogy a logatomok esetében a leggyakrabban tévesztett hang az [y] volt, míg az [ɛ]-t egyszer sem azonosították tévesen, ami az jelenti, hogy a spontánbeszéd-vizsgálatok alapján (Gósy 2004) legritkábban előforduló hang okozza a legtöbb téves észleletet, míg a leggyakoribb a legkevesebbet. Valószínűsíthető tehát, hogy a mindennapi kommunikációban gyakrabban előforduló hangokat könnyebben azonosítja a hallgató, mint a ritkábban szereplőket.

Mindezek alapján azonban a suttogott ejtés okozta módosulás az akusztikai kulcsokban nem okozhat jelentős nehézségeket a magánhangzók azonosításában a mindennapi kommunikáció során.

Irodalom

- Assmann, Peter F. – Nearey, Terrance M. é. n. Modelling the perception of frequency-shifted vowels. http://www.utdallas.edu/~assmann/ICSLP02_paper.pdf. A letöltés ideje: 2006. február.
- Gósy Mária 1984. *Hangtani és szótani vizsgálatok három éves gyerekek nyelvében*. Nyelvtudományi Értekezések 102. Akadémiai Kiadó, Budapest.
- Gósy Mária 2002. Beszédképzés zöngé nélkül. *Beszédkutatás* 2002. 18–37.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Grácz Tekla Etelka 2005. Suttogott mássalhangzók zöngésségének észlelése. *Beszédkutatás* 2005. 78–91.
- Hillenbrand, James M. – Houde, Robert 2002. A narrow band pattern-matching model of vowel perception. <http://homepages.wmich.edu/~hillenbr/Papers/NarrowBandModel-JASA.pdf>. A letöltés ideje: 2006. február.
- Horváth Viktória 2005. A magánhangzók nazalizációjáról. *Beszédkutatás* 2005. 51–63.
- Ladefoged, Peter 2005. *Vowels and consonants: An introduction to the sounds of languages*. Second edition. Blackwell Publishing, Oxford.
- Morris, Robert W. – Clements, Mark A. 2002. Reconstruction of speech from whispers. *Medical Engineering and Physics* 24. 515–520.
- Stokes, Michel A. é. n. The MAS Model of Vowel Perception and Production. <http://home.indy.net/~masmodel/masmodel.pdf>. A letöltés ideje: 2006. január.
- Strange, Winifred 1999. Perception of vowels: Dynamic constancy. In Pickett, James M. (ed.): *The acoustics of speech communication. Fundamentals, speech perception theory, and technology*. Allyn and Bacon, Boston–London–Toronto–Sidney–Tokyo–Singapore, 153–165.

A MEGAKADÁSJELENSÉGEK AKUSZTIKAI ÉS PERCEPCIÓS SAJÁTOSSÁGAI

Bóna Judit

Bevezetés

Beszédünk mintegy egyharmadát megakadásjelenségek teszik ki, amelyek a beszédtervezés és a beszédkivitelezés összehangolatlanságából származnak (Gósy 2005: 107). A megakadások többsége a beszélő bizonytalanságából származó néma vagy kitöltött szünet, azaz hezitálás.

A megakadásoknak két nagy típusa van: a bizonytalanságból adódóak (bizonyos néma szünetek, hezitálás, töltelékszó, ismétlés) és a téves kivitelezések (Gósy 2002). A hiba típusú megakadásjelenségek aránya jóval kisebb, bár mértékük beszélőnként erősen változhat (Gósy 2002; Huszár 2005). Megjelenésük függ a beszélő személyiségétől, a beszélők életkorától (Menyhárt 2003), a beszédtempótól (Bóna 2006) és számos más tényezőtől is.

A téves kivitelezések a beszédprodukciónak bármely szintjén létrejöhetnek, ezért nagyon sokfélék lehetnek. Az első magyar „Nyelvbottlás”-korpusz (Gósy et al. szerk. 2004, 2005) a következő megakadásokat különíti el: freudi elszólás; morfológiai, szintaktikai hiba, ezen belül újraindítás nélküli morfológiai változtatás és ismétlés toldalékjavítással; kontamináció; téves szótalálás, ezen belül téves szókezdés, illetve malapropizmus; „nyelvem hegyén van” jelenség; változtatás; újraindítás, ezen belül újraindítás változtatással; perszeveráció; anticipáció; metatézis; egyszerű nyelvbottlás (betoldás, csere, kiesés). Ezek az újraindítást lezámítva mind a hiba típusú megakadások körébe tartoznak.

A megakadásjelenségeket kétféleképpen szokás gyűjteni, a gyűjtési módtól függően kétféle megakadaskorpuszt különböztethetünk meg: az *on-line* vagy *valós idejű* jegyzetelt, illetve az *off-line* vagy rögzített korpuszokat. (Az *on-line* és *off-line* terminusokat külföldi szakirodalom alapján Huszár Ágnes használta először a magyar szakirodalomban [2005]; a *valós idejű* és *rögzített* elnevezések Markó Alexandrától [2006] származnak.) A *valós idejű* gyűjtéskor a gyűjtő a hétköznapi beszédhelyzetekben hallott megakadásokat jegyzeteli, vagyis ekkor a saját percepciós mechanizmusa által kiszűrt jelenségeket veszi észre. A legnagyobb és elsőként publikált magyar nyelvű jegyzetelt korpuszt 2004-ben, majd folytatónak 2005-ben jelentették meg, ez a már említett „Nyelvbottlás”-korpusz. A rögzített korpuszok esetén különféle spontánbeszéd-felvételekből gyűjti ki a kutató a beszédben előforduló megakadásokat. Magyar nyelven publikáltak adatokat monologikus spontánbeszéd-felvételekről (Gósy 2003); térképmódszerrel rögzített dialógusokról (Horváth 2004); társasjáték során rögzített hangfelvétel-

ről (Szabó 2004); többféle interjúhelyzetben készült spontánbeszéd-felvételről (Markó 2004); valamint a médiaszereplők megakadásairól (Gyarmathy 2005; Kaló 2005).

A kísérletek szerint a hallgatónak nyelvtől függetlenül szüksége van a megakadásokra, hiszen azok időt adnak neki az elhangzottak észlelésére, megértésére (Gósy 2000; Fox Tree 2001; Susca–Healey 2002). A beszédproduktív megakadások jó részét a hallgató észre sem veszi, illetve a perzeptív mechanizmusa gyorsan korrigálja azt. Egy angol nyelvű kísérletben a résztvevőknek egyszeri meghallgatás után tesztalapon kellett jelölniük az elhangzó megakadásokat. Ezután többször is meghallgathatták a szöveget, és újra meg kellett találniuk a megakadásokat. A többszöri meghallgatás után a kísérleti személyek háromszor annyi megakadást találtak meg, mint egyetlen meghallgatás után (Ferber 1991).

Magyar nyelven először a szünetek perzeptívájával kapcsolatban születtek tanulmányok (Kassai 1988; Kassai–Fagyal 1996; Menyhárt 1998; Gósy 2000). A néma szünetnek mint megakadásnak az azonosítása nem egyszerű feladat, hiszen nemritkán nehezen dönthető el, hogy pusztán levegővétel a funkciója, netán a beszélő szándékosan idézte elő az értelmi tagolás érdekében, vagy a beszédtervezési folyamat során például a mentális lexikonban való kereséshez biztosítja az időt. Gósy Mária kísérletében (2000) függetlenül a funkciótól a szünetek észlelésének sajátosságait, illetve a szöveg szünetezésétől függően a beszédmegértést vizsgálta. Az eredmények azt mutatják, hogy függetlenül az aktuális tartalomtól, tempótól vagy feladattól, az emberek a beszédsszünetek mintegy kétharmadát észlelik; és a szünet időtartama és észlelése szoros összefüggést mutat.

A szünettől különböző megakadásjelenségek perzeptívájával foglalkozó első magyar nyelvű kísérletet Gósy Máriával végeztük 2005-ben, amelyben 10 férfi és 10 nő vett részt. A tesztanyagban kilenc kategóriából összesen 45 megakadásjelenség szerepelt, a kísérlethez nem választottunk olyan jelenséget, ahol a szándékolt közlés rendszerint csak az adott beszédsszituációból következtethető ki (freudi elszólás), illetve olyat, amelyben a szándékolt közlés elhangzik a beszélőtől („nyelvem hegyén van” jelenség javítása, változtatás, újraindítás). A kísérleti személyek feladata az volt, hogy a kísérletvezető által felolvasott megakadásjelenséget javítsák ki. Az eredmények alapján megállapítottuk, hogy a hallgató képes az elhangzott közlések hibáinak javítására; ez a javítás azonban nem tökéletes; időt vesz igénybe, amelynek tartama függ a megakadás típusától, a típuson belül is az adott példától és kontextustól (például a szógyakoriságtól), illetőleg a hallgatótól. Emellett minél magasabb tervezési szinten következik be a produktív hiba, annál bizonytalanabban működik a hallgató korrekciós stratégiája, és annál hosszabb időt vesz igénybe a javítás. Jellemzőek az egyéni különbségek, és a férfiak lassabbak a megoldás során (Gósy–Bóna 2006).

Markó (2006) valós idejű jegyzetelt és rögzített korpuszok segítségével vizsgálta a megakadásjelenségek hatását a beszédészlelésre. A kétféleképpen gyűjtött korpuszok adatainak összevetéséből kitűnik, hogy a valós idejű és a rögzített korpuszok között jellegzetes eltérések vannak. A spontán beszédben előforduló

újraindítás sokkal gyakoribb jelenség, mint ahányszor hallgatóként felfigyelünk rá, hasonlóan a változtatáshoz. Szintén ritkán tűnik fel a hallgató számára a grammatikai összehangolatlanság. Ezzel szemben a hallgatói percepcióban nagyobb arányban tűnnek fel a kontaminációk, a téves szótalálások, a sorrendiségi hibák és az egyszerű nyelvbontások.

Gyarmathy (2006) tanulmányában a beszédprodukciós és beszédpercepcióss folyamatok összefüggéseit vizsgálta a megakadásjelenségek által. A kísérletben 2 férfi és 3 nő vett részt, akiktől a kísérletvezető mintegy 20-20 perces spontán beszéd-felvételt rögzített, majd a kísérleti személyeknek egy rádiófelvétel meghallgatása során jelezniük kellett, amikor megakadásokat észleltek. Ezt a kísérletvezető jegyezte fel a tesztlapra, vagyis a vizsgált személyeket nem segítette az olvasott szöveg a percepció során. A kísérleti személyek a tesztszöveg megakadásainak csupán egy részét vették észre, egyénenként eltérő mértékben. Az eredmények azt mutatják, hogy a beszélők számára a saját beszédprodukciójukat kevésbé jellemző megakadások bizonyulnak feltűnőbbnek. Mind a beszédprodukcióban, mind a beszédpercepcióban a bizonytalanságból adódó jelenségek fordulnak elő a legnagyobb arányban. Az észlelést az egyéni sajátosságokon túl befolyásolja az adott megakadásjelenség előfordulási gyakorisága, illetve a megakadás beszélő általi korrekciója, független azonban az egyén nemétől.

A kísérletek eredményei tehát egyöntetűen arra világítanak rá, hogy a beszédprodukciós megakadások nagy részét a hallgató észre sem veszi, illetőleg igen gyorsan korrigálja. Jelen vizsgálatunk célja az, hogy meghatározzuk, milyen összefüggések vannak az egyes hiba típusú megakadások és a szünettartás között, illetve milyen tényezők befolyásolják a különféle megakadásjelenségek percepcióját. Vizsgálatunk újdonsága az is, hogy a valós idejű és rögzített korpuszok adatainak összehasonlítását ugyanazon beszédfelvétel alapján tesszük meg.

Hipotéziseink szerint 1. a megakadásjelenségeket gyakran szünetek követik vagy előzik meg, különösen a bizonytalanságból adódóakat; 2. a magasabb tervezési szinteken bekövetkező megakadásokra jobban felfigyel a hallgató, hiszen nehezebb javítani őket; 3. a rögzített gyűjtésben több bizonytalanságból eredő megakadásjelenség lesz; 4. a hallgató jobban felfigyel a megakadásokra, ha közben olvashatja a hibátlan szöveget.

Anyag és módszer, kísérleti személyek

Vizsgálatunkhoz egy férfi beszélővel mintegy 7,5 perces spontán narratívát rögzítettünk digitális formában. A beszélő idegenvezetőként dolgozik, külföldieknek vezet túrákat Budapesten, illetve Egerben. Ezért arra kértük, hogy az utazási iroda programválasztékáról, az egyes túratípusokról meséljen a felvételen. A spontán beszédet lejegyeztük, majd a Praat 4.2-vel meghatároztuk benne a szünetek számát és a hosszúságát.

A szakirodalomban a szünet minimális hosszának többféle értéket, 30–150 ms-ot szoktak megadni (vö. Gósy 2000). Méréseinkben minden olyan jel nélküli szakaszt szünetnek tekintettünk és megmértünk, amely független volt a

beszédhangok képzésétől, és hallás alapján is egyértelműen érzékelhető volt. Így a legrövidebb szünettartam 100 ms-nál rövidebb, 70 ms lett.

Ezután kigyűjtöttük, majd a „Nyelvbottlás”-korpusz kategóriáinak megfelelően csoportosítottuk a beszédben talált megakadásjelenségeket. Bizonyos megakadásokat a különböző szerzők különféleképpen osztályoznak, ilyenek a szünet a szóban és a változtatás. Saját osztályozásunkban a változtatást a hiba típusúakhoz, míg a szóban tartott szünetet a bizonytalanságból adódóakhoz soroltuk. Mivel a néma szünetről nehéz eldönteni, hogy megakadásjelenség vagy levegővételezt biztosít, netán az értelmi tagolás miatt a beszélő szándékosan tartotta, a néma szüneteket a megakadásjelenségektől külön kezeltük. Megvizsgáltuk, van-e összefüggés az egyes megakadástípusok és a szünettartás sajátosságai között.

A beszédprodukció elemzését követően egy percepcióis tesztet végeztünk el. A tesztben 20 egyetemista vett részt, férfiak és nők vegyesen. A kísérleti személyek háromszor hallgatták meg a spontánbeszéd-felvételt, s az volt a feladatuk, hogy 1. az első meghallgatás után egy üres papírra jegyezzék le az észlelt megakadásokat; 2. a második meghallgatás során egy tesztlapot kaptak, amelyen az elhangzó szöveg hibátlan mondatokba átírva szerepelt, s ebben kellett jelölniük a megakadásjelenségeket; 3. a harmadik meghallgatás közben pedig azt kértük, jelöljék ugyanezen lapon a szüneteket is. Mivel a tesztlap a megakadások nélküli, javított szöveget tartalmazta, az olvasott és a hallott szöveg különbözősége segíthette a kísérlet résztvevőit. A hallgatókkal a teszt előtt röviden ismertettük a megakadásjelenségek és a szünet definícióját, a megakadások altípusait, és példákkal is illusztráltuk őket.

Mivel a megakadásjelenségek gyűjtése a tartalomról magára a beszédfolyamatra irányítja a hallgató figyelmét, kíváncsiak voltunk, ez mennyiben befolyásolja a szövegértést. (Annál is inkább szerettük volna ezt tudni, mivel a saját hallgatóink az egyetemi szemináriumok alatt is gyűjtik ezeket a jelenségeket.) Ennek megválaszolásához a második meghallgatás után megértést ellenőrző kérdéseket tettünk fel a kísérleti személyeknek, ezek a megértés mellett a memóriát is tesztelték. A kérdések és a válaszok a következők voltak: 1. *Hányféle túrát tart az utazási iroda? Sorolja fel őket!* (gyalog-, biciklis, kocsma-, éjszakai, gasztro-, „kommunista”, egri túra); 2. *Melyik városokban szerveznek túrákat?* (Budapest, Eger); 3. *Melyik túrát tartja a legnépszerűbbnek a beszélő?* (abszolút gyalogtúra); 4. *Milyen túrát nem vezetett még a beszélő?* (egri, gasztrotúra); 5. *Honnan érkeznek főként az iroda vendégei?* (Amerika, Anglia).

Az adatokat megvizsgáltuk abból a szempontból, hogy mely megakadásjelenségeket vették leginkább észre a hallgatók, illetve befolyásolja-e a szünettartás az egyéb típusú megakadásjelenségek észlelését. Mivel a rögzített szövegből kigyűjtött megakadások rögzített korpuszt alkotnak, az első meghallgatáskor lejegyzett megakadások pedig ugyanazon szöveg jegyzetelt gyűjtésű korpuszát adják, így lehetőségünk nyílt összehasonlítani a rögzített és a valós idejű adatgyűjtés jellemzőit is.

Eredmények

A beszédprodukción vizsgálata

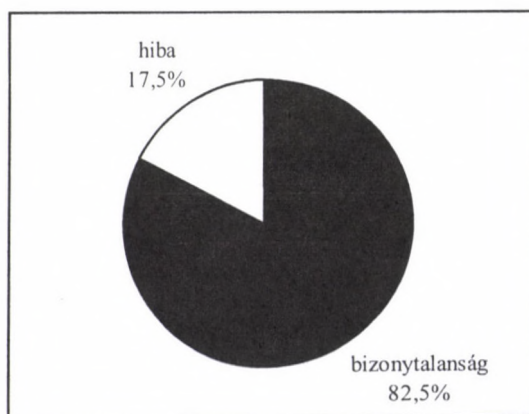
Az 1. táblázatban a spontán beszéd temporális adatait olvashatjuk. A beszélő közepes sebességgel beszélt, viszonylag sok szünetet tartott, és keveset hezitált.

1. táblázat: A spontán narratíva temporális adatai

| Temporális sajátosságok | Adatok |
|---|-------------------------|
| Beszékidő | 451 420 ms (7,524 perc) |
| Jelidő | 311 502 ms (5,192 perc) |
| Szünetidő | 139 918 ms (2,332 perc) |
| A szünetek aránya a teljes beszékidőhöz képest | 31% |
| A néma szünetek aránya az összes szünethez képest | 90,4% |
| A szünetek száma percenként | 32,3 db |
| Beszédtempó | 9,90 hang/s |
| Artikulációs tempó | 14,35 hang/s |

A beszélő a 7,5 perces monológja során 209 néma és 34 kitöltött szünetet tartott. A néma szünetek átlagos tartama 605 ms volt, a legrövidebb általunk mért szünet 70 ms, a leghosszabb 2548 ms volt. A beszélő 27-szer tartott 1 másodpercnél hosszabb szünetet.

A szövegben összesen 131 megakadásjelenséget találtunk – ez körülbelül 17 megakadást jelent percenként, ami jóval több, mint a korábbi méréseinkben kapott adatok átlaga (Bóna 2006) –, ezek több mint háromnegyede bizonytalanságból adódó, és körülbelül egynegyede hiba típusú volt. A 1. ábrán ezeket az arányokat láthatjuk.

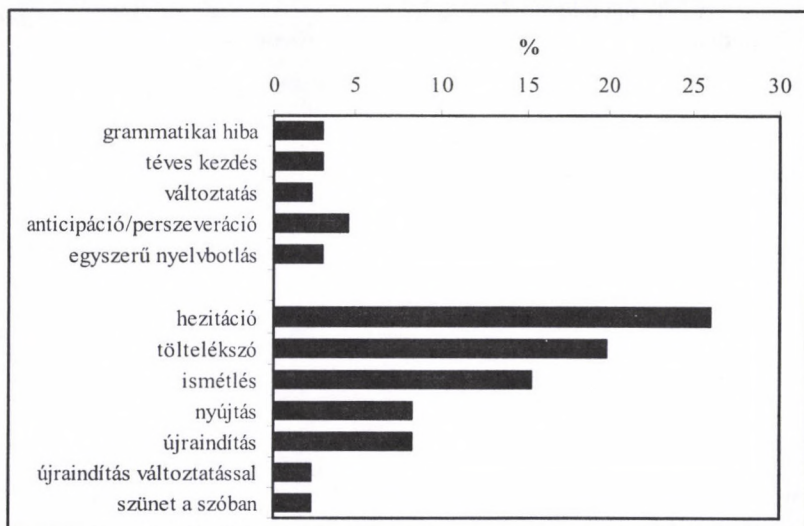


1. ábra

A bizonytalanságból adódó és a hiba típusú megakadások aránya a vizsgált spontán beszédben

A vizsgált spontán beszédben nem fordult elő freudi elszólás, kontamináció, „nyelvem hegyén van” jelenség, illetve metatézis.

A hiba típusú megakadások között a sorrendiségi hibák voltak a leggyakoribbak, de alig többször fordultak elő, mint a többi hiba típusú megakadás. A bizonytalanságból adódó megakadások között a hezitálásból találtuk a legtöbbet, de igen gyakoriak voltak a töltelékszavak és az ismétlések is (2. ábra).



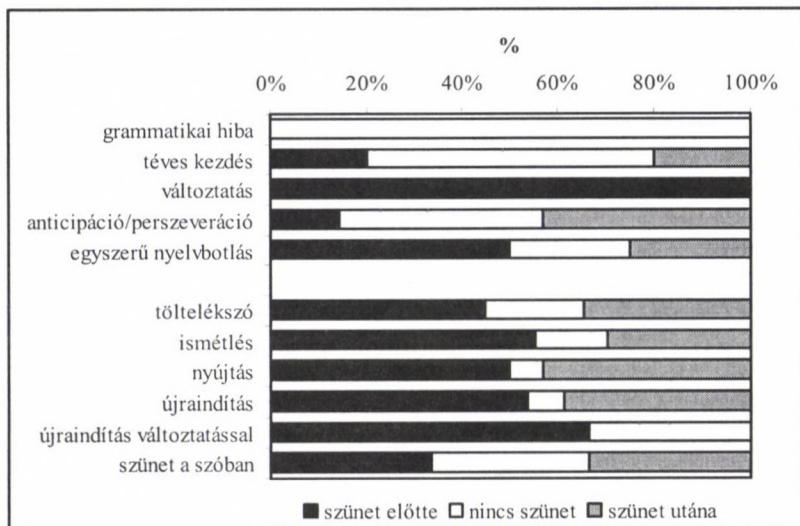
2. ábra

Az egyes megakadástípusok aránya a korpuszban

Hezitálások, más típusú megakadások és néma szünetek gyakran jelentek meg együtt. Vagyis a legtöbb megakadás előtt és/vagy után a beszélő szünetet tartott. A néma és kitöltött szünetek, illetve az egyes megakadástípusok összefüggéseit a 3. ábráról olvashatjuk le.

A vizsgált szövegben a grammatikai hibák esetén nincs szünet; a változtatás esetén mindig van, mielőtt elkezdené a beszélő az új egységet; az újraindítás és a nyújtás esetén szinte mindig találtunk szünetet. Az újraindítás, újraindítás változtatással és az ismétlés esetén nagyon gyakran az egész közlés előtt, majd a megismételt/újraindított elem előtt is van szünet, gyakran utána is. Ha a megakadások előtt és után levő szünetek átlaghosszúságát kiszámítjuk, megállapítható, hogy a megakadások előtti szünetek átlagosan hosszabbak, mint a megakadások után megjelenők. Például: *legalábbis ez a* (651 ms) *ez a* (281 ms) *kör; megnézzük a* (979 ms) *aa* (401 ms) *különböző* (1096 ms) *őö* (389 ms) *hát; a* (239 ms) *Vár* (130 ms) *Várnegyedben*.

Nemcsak hiba típusú megakadások és szünetek álltak egymás mellett, hanem gyakran egész megakadásjelenség-sorozatot találtunk. Ilyenek voltak például: *hanem mindenféle őő* (= 439 ms)¹ *olyan* (815 ms) *őő* (= 385 ms) (768 ms) *hát; így* (658 ms) *őő* (= 518 ms) (1448 ms) *hát még v nem említettem; a* (1243 ms) *a túra* (707 ms) *i* (553 ms) *idő* (73 ms) *ideje is* (491 ms) *időtartama*.



3. ábra

A szünettartás és a megakadások összefüggései

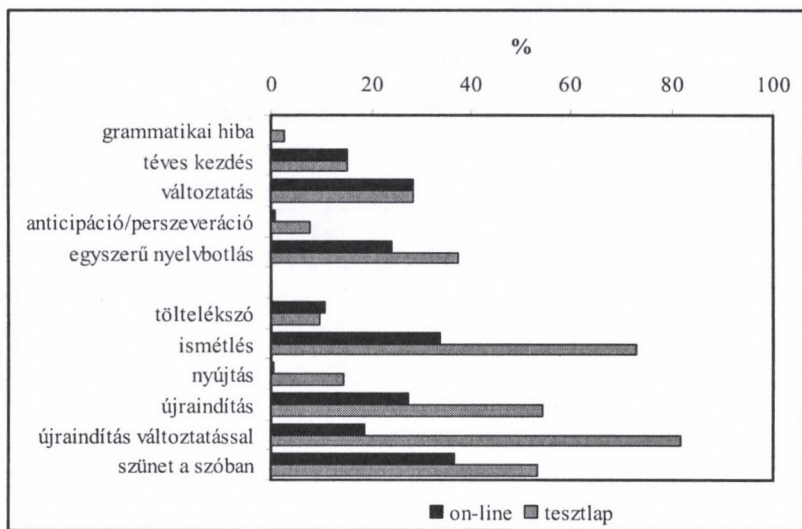
A percepcióss teszt eredményei

Az első meghallgatás közben, amikor üres lapra, az on-line vagy valós idejű gyűjtés módszere szerint írták le a kísérleti személyek a megakadásokat, várhatóan csak igen kis hányadát (17,5%-át) sikerült az összes megakadásnak lejegyezniük. A második meghallgatás során, amikor már a tesztalapon szereplő, hibátlan szövegben kellett jelölniük a megakadásokat, ez az arány 33,8%-ra változott. A 4. ábrán a két meghallgatás során helyesen felismert megakadások arányait láthatjuk. A hallgatók a megakadások 5,7%-át hibásan észlelték, vagyis megakadást jelöltek olyan helyeken, ahol az nem fordult elő.

A kísérletben részt vevők a legkevésbé a grammatikai hibákra és a nyújtásokra figyeltek fel (a valós idejű gyűjtéskor ezeket egyáltalán nem jegyezték le), és alig tűntek fel számukra az anticipációk és a perszeverációk is. A valós idejű

1. A zárójelen belüli egyenlőségjel arra utal, hogy az érték a kitöltött a szünet (őő, mm stb.) időtartama, míg az egyenlőségjel nélkül álló adatok az adott helyen megjelenő néma szünetek időtartamát mutatják.

gyűjtés során sikerült körülbelül az egyharmadát lejegyezniük a beszédben található ismétléseknek és szavakban tartott szüneteknek, illetve a változtatásoknak is. A tesztalapon való jelölés egyrészt azért bizonyulhatott könnyebbnek, mert ekkor másodszorra hallhatták a szöveget, másrészt olvashatták. Ez az olvasásból származó könnyebbség azonban nem minden esetben érvényesült, hiszen a töltélszavak felismerésére nem volt hatással: ezeket mindegyik gyűjtés során csak nagyon kis arányban sikerült megtalálniuk a kísérleti személyeknek. Ez az eredmény is alátámasztja azt a feltételezést, hogy a töltélszavakra kevésbé figyelünk fel, mivel nem akasztják meg a beszéd folyamatosságát, nem gátolják a beszédfeldolgozási mechanizmust. Valószínűleg a vizuális és az akusztikai inger különbözősége segítette a hallgatókat az újraindítások, az újraindítás változtatással, illetve az ismétlések sokkal nagyobb arányú felismerésében a tesztalapos jelölés során. Ekkor több mint 80%-ot sikerült a hallgatóknak jelölni az újraindítás változtatással jelenségből. Például azt a megakadást, hogy *személyesebb, mint egy busztúra, ugye* (635 ms) *mint egy buszos túra* a valós idejű gyűjtés során a hallgatók mindössze 25%-a, míg a tesztalapos jelölés során 70%-uk észlelte.

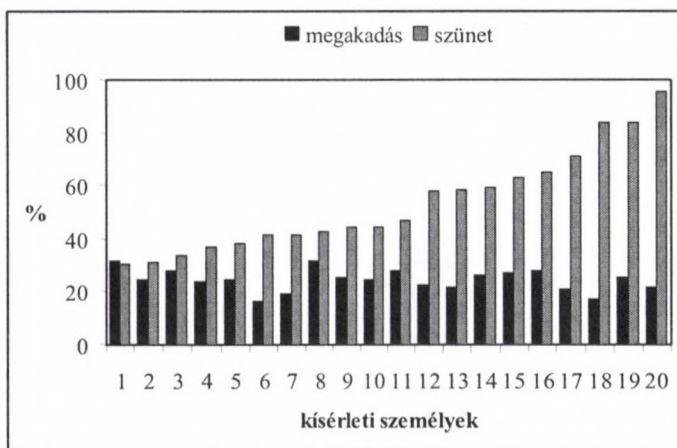


4. ábra

A helyesen felismert megakadások a valós idejű/on-line gyűjtés és a tesztalapon való jelölés során (100% = az összes megakadásjelenség)

A szünetészlelési kísérletben a hallgatók a szünetek mintegy 54%-át észlelték helyesen, az összes észlelt szünet 10%-át hibásan. A legjobban teljesítő személy a szünetek 96%-át, a legrosszabbul teljesítő 31%-ot észlelt helyesen. A kísérleti személyek általában a hezitálásokat és az 500 ms-nál hosszabb néma szüneteket

észlelték jól. A szünetészlelés és a megakadások észlelése között nem találtunk összefüggést, hiszen a szüneteket legjobban észlelő személy az átlag alatt teljesített a megakadások észlelésében, míg a legjobb megakadásészlelő személy átlagon alul ismerte fel a szüneteket. Mindkét feladat eredményeiben nagy egyéni különbségek vannak (5. ábra).



5. ábra

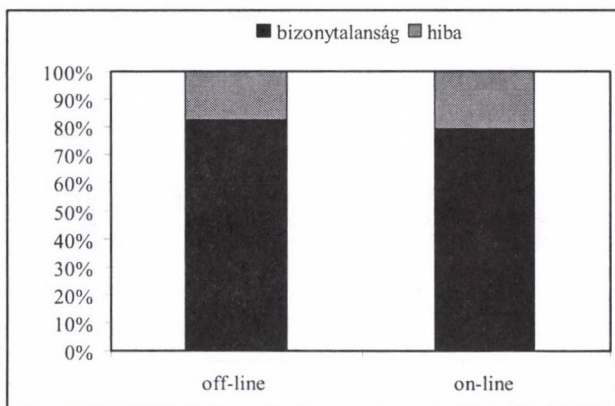
A szünetek és a megakadások észlelésének összefüggései

Várakozásunkkal ellentétben a szünetek nem feltétlenül segítettek a megakadások észlelését. A két meghallgatás eredményei azt mutatják azonban, hogy sokkal kevésbé sikeres a megakadás észlelése, ha nem áll előtte és/vagy utána szünet. A szünet hosszúsága vagy kitöltött/kitöltetlen volta azonban nem befolyásolta az észlelést, valószínűleg más beszédtenyezőkben kell keresnünk a megakadások különböző percepciójának az okát. Például a legtöbb hallgató által felismert megakadások a következők voltak: *itthonai itthoni mm* (= 460 ms) (134 ms) *irodák közül; igyekszünk* (285 ms) *bele* (494 ms) *beleszorítani; hogy mi* (280 ms) *mit* (321 ms) *érdemes erről tudni; az benne a jó, hogy* (90 ms) *hogy ő* (= 466 ms) *sokkal személyesebb*. A legkevesebben – azokon a megakadásokon kívül, amelyeket nem előz meg vagy nem követ szünet – az alábbi megakadásokat ismerték fel: *jellemző lehet a magyar* (861 ms) *étkezés* (487 ms) *étkezés-re* (1717 ms); *nem említettem a* (750 ms) *az* (123 ms) *éjszakai gyaloglást; s utána* (2548 ms) *azután még* (295 ms). A példákon megfigyelhetjük, hogy a kevesebb hallgató által felismert megakadások mellett hosszabb szünetek is állhatnak, mint a sokak által észlelt megakadások mellett.

A kétféle gyűjtésű korpusz adatainak összehasonlítása alátámasztja a szakirodalmi adatokat. A tendencia szerint hallgatóként kisebb arányban figyelünk fel a

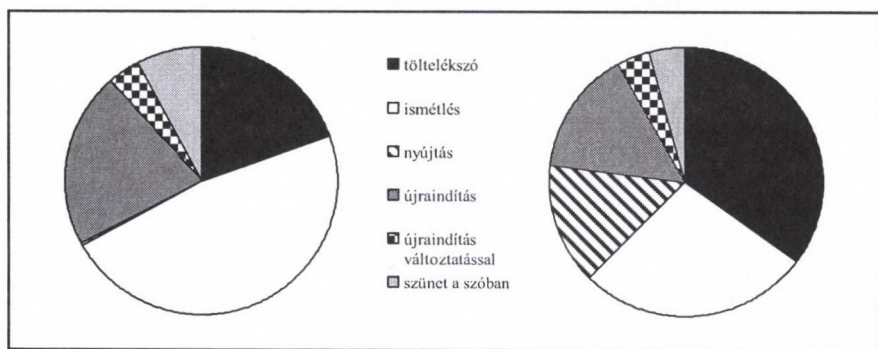
bizonytalanságból adódó megakadásokra, hiszen funkciójuk az, hogy időt biztosítsanak a feldolgozás számára; ugyanakkor a valós idejű gyűjtésben növekszik a hiba típusú megakadások aránya (6. ábra).

A két nagy kategórián belül is jellegzetes különbségek vannak az egyes megakadástípusok megjelenései között a korpuszokban (7. ábra). Amíg a valós idejű gyűjtésben főként ismétlések szerepelnek, addig a rögzített korpuszban a töltelékszavakból találjuk a legtöbbet. Ebben a korpuszban jóval több a nyújtás felismerése is.



6. ábra

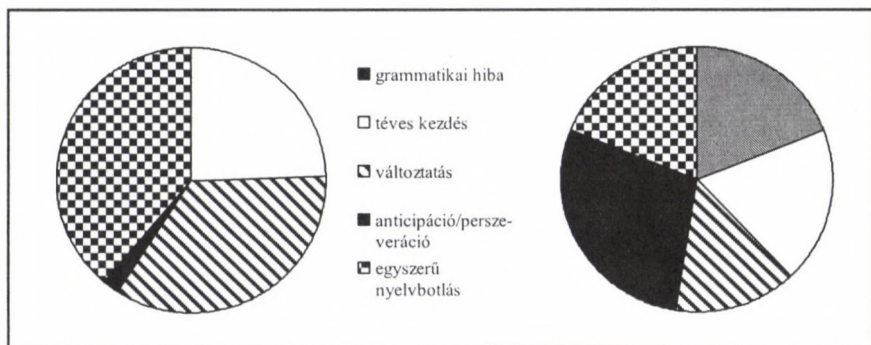
Bizonytalanságból származó és hiba típusú megakadások a rögzített/off-line és a valós idejű/on-line korpuszban



7. ábra

A bizonytalanságból származó megakadástípusok a valós idejű (bal oldalon) és a rögzített (jobb oldalon) korpuszban

A hiba típusú megakadások között is jellegzetes eltéréseket találunk korpusztól függően (8. ábra).



8. ábra

A hiba típusú megakadások a valós idejű (bal oldalon) és a rögzített (jobb oldalon) korpuszban

Adataink azonban csak részben egyeznek Markó (2006) adataival, hiszen a valós idejű korpuszban több újraindítást találunk, kevesebb a téves kezdés, illetve a sorrendiségi hiba, mint a rögzített korpuszban. (Az újraindítás és a sorrendiségi hibák esetén viszonylag nagyok a különbségek Markó adatai és a saját adataink között, igaz, a vizsgált megakadások száma között is. Újraindítás Markó vizsgálatában 6,68%-ban fordult elő a valós idejű, és 22,57%-ban a rögzített korpuszban, a jelen vizsgálatban 21,3% a valós idejű és 10,7% a rögzített korpuszban. A sorrendiségi hibák Markónál 32,55%-ot tesznek ki a valós idejű és 15,78%-ot a rögzített korpuszban; saját anyagunkban ez az arány 0,3% a valós idejű gyűjtéskor és 4,58% a rögzített korpusz alapján. A téves kezdés esetén alig találtunk különbséget a két korpusz között: 3,6% a valós idejű, 3,05% a rögzített korpuszban; míg Markó tanulmányában 4,84% a valós idejű és 13,03% a rögzített korpuszban.) Az adatok a grammatikai hibákat és az egyszerű nyelvbotlást illetően viszont megegyeznek Markó adataival.

A szövegértést vizsgáló kérdésekre – hipotézisünknek megfelelően – 63,4%-ban érkeztek jó válaszok. A legjobban a 2. és az 5. kérdést tudták megválaszolni a kísérleti személyek, ezekre 90%-ban, illetve 85%-ban adtak jó válaszokat. A legnehezebb a 3. kérdés volt, pedig a válasz többször elhangzott, és erről mesélt a legtöbbet a beszélő. A szövegértési teszt eredménye arra hívja fel a figyelmet, hogy a megakadások gyűjtése közben a gyűjtő nem tud igazán az elhangzó szöveg tartalmára figyelni, még akkor sem, ha kétszer hallgathatja meg a szöveget, és eközben egyszer olvashatja is.

Következtetések

Hipotéziseink közül tehát beigazolódott, hogy a megakadásjelenségek gyakran társulnak szünettartással. Megfigyeltük azonban, hogy a szünet hosszúsága vagy kitöltött/kitöltetlen volta nem befolyásolta az észlelést. Ennek alapján valószínűsíthetjük, hogy az akusztikai környezetnél (itt: szüneteknél) nagyobb hatása van az észlelésre a megakadás típusának, illetve annak, hogy az hol jelenik meg a mondatban/szövegben.

Kísérletünk alátámasztotta azt a szakirodalmi megállapítást, hogy a valós idejű gyűjtés kevesebb megakadás észrevételét teszi lehetővé, mivel ez a hallgató egyszeri percepcióján alapszik; és a valós idejű korpuszokban tendenciaszerűen valamivel kevesebb a bizonytalanságból adódó megakadás, mint a rögzített korpuszban.

A hallgatók többszöri meghallgatás után, illetve a szöveg olvasásakor több megakadásjelenséget vettek észre, de még ekkor is több mint a megakadások fele rejtve maradt számukra. Az eredmények rávilágítanak arra is, hogy a hallott szöveg olvasása nem feltétlenül segíti a megakadások felismerését. A megakadásjelenségek gyűjtése közben a gyűjtő nem tud tökéletesen az elhangzottakra figyelni.

Irodalom

- Bóna Judit 2006. Tudunk-e változtatni spontán beszédünk tempóján? In Mártonfi Attila – Papp Kornélia – Slíz Mariann (szerk.): *101 írás Pusztai Ferenc tiszteletére*. Argumentum Kiadó, Budapest, 560–566.
- Ferber, Rosa 1991. Slip of the tongue or slip of the ear? On the perception and transcription of naturalistic slips of the tongue. *Journal of Psycholinguistic Research* 20. 105–22.
- Fox Tree, Jean E. 2001. Listeners' uses of *um* and *uh* in speech comprehension. *Memory and Language* 29. 320–326.
- Gósy Mária 2000. A beszédszünetek kettős funkciója. *Beszéd kutatás* 2000. 1–15.
- Gósy Mária 2002. A megakadásjelenségek eredete a beszédprodukción tervezési folyamatban. *Magyar Nyelvőr* 126. 192–204.
- Gósy Mária 2003. A spontán beszédben előforduló megakadásjelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária – Bóna Judit 2006. A megakadásjelenségek javítása a beszédmegértésben. *Magyar Nyelvőr* 130. 33–49.
- Gósy Mária et al. (szerk.) 2004. „Nyelvbotlás”-korpusz. *Beszéd kutatás* 2004. 19–186.
- Gósy Mária et al. (szerk.) 2005. „Nyelvbotlás”-korpusz – 2. rész. *Beszéd kutatás* 2005. 145–173.
- Gyarmathy Dorottya 2005. *Különböző beszélők egyéni megakadásjelenségeinek sajátosságai*. Szakdolgozat. ELTE, Budapest.
- Gyarmathy Dorottya 2006. A beszédpercepció és beszédprodukción folyamat összefüggései a megakadásjelenségek tükrében. Megjelenés alatt.

- Horváth Viktória 2004. Megakadásjelenségek a párbeszédekben. *Beszéd kutatás* 2004. 187–199.
- Huszár Ágnes 2005. *A gondolatól a szóig. A beszéd folyamata a nyelvbotlások tükrében.* Tinta Könyvkiadó, Budapest.
- Kaló Zsuzsa 2005. *A nyilvános spontán beszéd megakadásjelenségei és a médiaszereplők szóaktiválási stratégiái.* Szakdolgozat. ELTE, Budapest.
- Kassai Ilona 1988. A szünet kérdésköre a szöveglejegyzésben. In Kontra Miklós (szerk.): *Beszélt nyelvi tanulmányok.* Linguistica, Series A, Studia et Dissertationes 1. MTA Nyelvtudományi Intézet, Budapest, 22–43.
- Kassai Ilona – Fagyal Zsuzsanna 1996. Hogyan észlelik a magyar beszéd szüneteit magyar és francia anyanyelvű hallgatók? *Magyar Nyelvőr* 120. 209–220.
- Markó Alexandra 2004. Megakadások vizsgálata különféle monologikus szövegekben. *Beszéd kutatás* 2004. 209–222.
- Markó Alexandra 2006. A megakadásjelenségek hatása a beszédészlelésre. *Alkalmazott Nyelvtudomány* VI/1–2. 103–118.
- Menyhárt Krisztina 1998. Nyelvi meghatározottság a beszédsszünetek észlelésében. *Beszéd kutatás* '98. 47–57.
- Menyhárt Krisztina 2003. A spontán beszéd megakadásjelenségei az életkor függvényében. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia a gyakorlatban.* Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 125–138.
- Susca, Mike – Healey, Charles E. 2002. Listener perceptions along a fluency-disfluency continuum: A phenomenological analysis. *Journal of Fluency Disorders* 27. 135–161.
- Szabó Eszter 2004. Megakadásjelenségek nyelvi játék közben. *Beszéd kutatás* 2004. 200–208.

A tanulmány az OTKA T046455 számú pályázata keretében jött létre.

A MEGGYŐZÉS A BESZÉDPRODUKCIÓBAN ÉS A BESZÉDPERCEPCIÓBAN

Bata Sarolta

Bevezetés

Napjaink népszerű retorikaszerzői – többek között – azzal az ígérettel ajánlják könyveiket, hogy hozzásegítik az olvasót a meggyőző kommunikációhoz: „Lehetővé teszi nézeteink világos és meggyőző kifejtését” (Hull 1997: 9, lásd még pl. Carnegie 1999). Ezek a könyvek igen kelendők, újabb és újabb kiadások születnek hasonló témakörökben, hiszen – ahogy Aczél Petra megfogalmazta: „korunk kulcsszava a meggyőzés” (Adamik–A. Jászó–Aczél 2004: 396).

A meggyőzni tudás vágya azonban egyáltalán nem új keletű, már az ókortól kezdve foglalkoztatja az embereket, hogy hogyan hathatnának igazán hallgatóságukra a beszédükkel. A hagyomány szerint a retorika kialakulásához is ez az igény vezetett, hiszen a Kr. e. 5. században szicíliai zsarnokok által elkobzott földeket kellett visszaszerezniük hatásos beszédekkel a lakosoknak (Barthes 1997). Arisztotelész szerint is az a retorika célja, hogy „feltárja minden téren a meggyőzés lehetséges módjait” (1999: 1355b). A beszéd a szerző *Rétorika* című műve alapján „három dologból áll össze: a beszélőből, amiről beszél, és akihez beszél” (i. h.). Mind a három tényezőt figyelembe kell venni a meggyőzés eléréséhez. A beszédet jól fel kell építeni, jó érveket kell felsorakoztatni, a tárgyhoz illő beszédfajtát kell alkalmazni, a megfelelő stílust és elrendezést kell megtalálni. A hallgatóra is tekintettel kell lenni; mindenképpen figyelembe veendő, hogy milyen összetételű a hallgatóság, a szónoknak pedig feddhetetlen jellemnek kell lennie. Azonban ez még mind nem elég. Már az ókori bölcs is úgy vélte: „nem elég azt tudni, hogy mit kell mondanunk, azt is tudni kell, hogyan adjuk elő” (Arisztotelész 1999: 1403b).

A többi ókori szerző is fontos szerepet tulajdonít az előadásmódnak, többek között Cornificius *A C. Herenniusnak ajánlott retoriká*-ban részletesen ismerteti, hogy az egyes beszédfajták során (társalgás, elbeszélés, tréfálgatás, buzdítás stb.) milyen hangerővel, hangmagassággal beszéljünk (Cornificius 2001). Cicero (1987) az előadás módjának szerepét méltatva így ír: a szónoknak „legeslegfőbb feladata, hogy arra fordítsa figyelmét, hogyan is mondja el” mondanivalóját. „Így aztán Demosthenes nem ok nélkül tulajdonítja az első, második és harmadik főszerepet az előadásmódnak” (220). Quintilianus korára pedig még fontosabbá vált az előadásmód, ő már egy egész fejezetet szentel ennek a témának, ahol részletezi, hogy az előadás módja hogyan legyen kifogástalan, világos és il-
lő (idézi: Adamik–A. Jászó–Aczél 2004: 118–121).

A beszéd művészetével foglalkozók ma is igen fontos szerepet szánnak könyveikben az előadásmódnak, ahol felhívják a figyelmet a megfelelő öltözködésre, térkihasználásra, hanggal való bánásmódra stb. (lásd Wacha 1994; Elekfi–Wacha é. n. [2003]). Ezek a könyvek azonban azt írják le, hogyan **kellene** beszélnünk ahhoz, hogy meggyőzőek legyünk, kísérletünkben pedig arra kerestük a választ, hogyan zajlik ez a valóságban, hogyan törekszenek a meggyőzésre a beszélők.

Kísérletsorozatunkban a beszédben megjelenő meggyőzést vizsgáltuk fonetikai szempontból, azon belül is a beszédprodukciónban és a beszédpercepcióban való megjelenését. A produkcióra vonatkozó kísérletben arra kerestük a választ, hogy az emberek hogyan teszik meggyőzőbbé beszédüket, ha a megszövegezésen nem változtathatnak, csak a szupraszegmentális tényezők állnak rendelkezésükre a cél eléréséhez. Hogyan befolyásolja tehát a meggyőzés szándéka a beszéd szupraszegmentális szerkezetét, megváltoznak-e a beszéd időviszonyai, a szünettartási jellegzetességek, a beszéddallam és a hangsúlyozás jellemzői? Ha találunk változást, vajon ezek tendenciaszerűek-e, beszélhetünk-e általános hangzásbeli „meggyőzési stratégiákról”? Percepciós kísérletünkben azt vizsgáltuk, vajon a meggyőzőnek szánt közlés valóban meggyőzi-e a hallgatóságot, azaz mennyire eredményesek meggyőzési stratégiáink; illetve a meggyőzés tekintetében melyek a legmegfelelőbb hangzási sajátosságok?

Anyag, módszer, kísérleti személyek

Produkciós kísérletünkben felolvasás során vizsgáltuk a meggyőzés szándékának hatását. Azért választottuk ezt a módszert, hogy valóban azt vizsgálhassuk, amire kíváncsiak vagyunk, azaz adatközlőink ne a szöveg megváltoztatásával (érveléssel, alakzatokkal, a beszéd felépítésének megváltoztatásával stb.) kívánják meggyőzni hallgatóságukat, hanem csakis a szupraszegmentális szerkezet megváltoztatásával. A felhasznált szöveg a következő volt: „Töltsön egy csodálatos hetet Rómában! Használja ki az Álom Utazási Iroda nagy lehetőségét! Mindössze három óra alatt odaérhet a varázslatos mediterrán városba. Szállás a város szívében egy elegáns szállodában teljes ellátással. Ókori romok, gyönyörű templomok, eredeti olasz ízek, ragyogó napsütés, barátságos emberek! Ne habozzon, még ma jelentkezzen álmai utazására!”

A kísérlet két részből állt: a kísérleti személyeknek az előbbi szöveget kellett felolvasniuk, elsőként minden instrukció nélkül. A második felolvasáshoz olyan szituációt próbáltunk teremteni, amelyben adatközlőink a meggyőzés szándékával lépnek fel, ezért az újrólvasás előtt egy írásbeli utasítást adtunk nekik: „Ön az Álom Utazási Iroda munkatársa. Ön mellett egy másik utazási iroda is ajánl római utazást. Olvassa fel úgy a hirdetést, hogy hallgatói elhiggyék, az Álom cég ajánlatánál nincs jobb!”

A vizsgálatban tíz nő és tíz férfi vett részt, átlagéletkoruk 29 év (18 és 60 év közöttiek), mindannyian magyar anyanyelvűek, budapesti értelmiségiek vagy egyetemisták.

Az adatokat a Praat beszédelemző program 4.3-as verziójával értékeltük ki. A kísérleti személyek két felolvasását – különböző szempontok alapján – páronként összehasonlítottuk egymással. Összevetettük a felolvasások időviszonyait mind a beszédtempó, mind az artikulációs tempó tekintetében. Vizsgáltuk a beszéd- és a szünetidő arányát, a szünetek átlagos előfordulási számát, a beszélők hangterjedelmét és a mondatok dallammenetét. Az adatokon statisztikai próbát végeztünk az InStat szoftver 3.0-s verziójával (párosított *t*-próba). A hangsúlyozás különbségeit független megfigyelővel végzett percepciósi teszteléssel vizsgáltuk.

Percepciósi kísérletünkben 103 adatközlő vett részt: 55 budapesti egyetemista, átlagéletkoruk 22 év (18 és 34 év közöttiek) és 48 budapesti középiskolás: 26 kilencedikes (15-16 évesek) és 22 hatodikos (12-13 évesek).

A percepciósi kísérlet anyagául azokat a felolvasáspárokat választottuk ki, amelyek jelentősen eltértek egymástól hangsúlyok, szünetek vagy időviszonyok tekintetében, így nyolc kísérleti személy két-két felolvasását. Véletlenszerű sorrendben játszottuk le a felvételeket, de arra ügyeltünk, hogy ugyanattól az adatközlőtől felvett két változat soha ne egymás után hangozzon el, mindig távolságot hagyunk köztük, nehogy a hallgatók az első változat alapján ítéljenek a második meghallgatásakor. Egyes beszélőktől az alapváltozatot hallották először a kísérletben részt vevők, míg másoktól a meggyőzőt, így kívántuk elkerülni, hogy a sorrend befolyásolja az eredményeket.

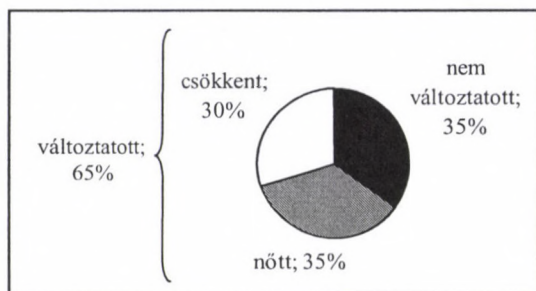
A percepciósi kísérletben részt vevők azt a feladatot kapták, hogy hallgassák meg a felvételeket, és értékeljék őket egyszeri meghallgatás után 1–5 pontszámmal annak alapján, hogy mennyire találják meggyőzőnek (1 = egyáltalán nem meggyőző, 5 = nagyon meggyőző). Lehetőségük volt a kiegészítő megjegyzések leírására is.

Eredmények

Produkciósi kísérlet

Az artikulációs tempó tekintetében nem vártunk számottevő eltérést egyazon beszélő két felolvasása között, hiszen az eddigi eredmények azt mutatják, hogy az artikulációs tempó jellemző az egyénre, és akaratlagosan igen kevésbé változtatható (vö. Gósy 2004 és Bóna 2005). Az eredmények igazolták hipotézisünket: az első felolvasásnál az átlagos artikulációs tempóérték 12,5 hang/s volt, a meggyőző felolvasásnál pedig 12,7 hang/s. A minimum- és a maximumértékek is szinte teljesen azonosak voltak (alap: 12,7 hang/s és 14,6 hang/s; meggyőző: 12,5 hang/s és 14,9 hang/s).

A beszédtempóértékeknél már nagyobb különbséget vártunk, hiszen ezt több tényező is befolyásolhatja: pl. a szünetek gyakorisága, hossza, illetve a kitöltött szünetek előfordulása (vö. Markó 2005: 72). Az eredmények alapján elmondhatjuk, hogy a kísérleti személyek 65%-a változtatott a beszédtempóján, azonban változtatás iránya nem tendenciaszerű: egyesek gyorsabban, mások pedig lassabban beszéltek, amikor meggyőzésre törekedtek, ezért összességében az eltérés nem szignifikáns (1. ábra).



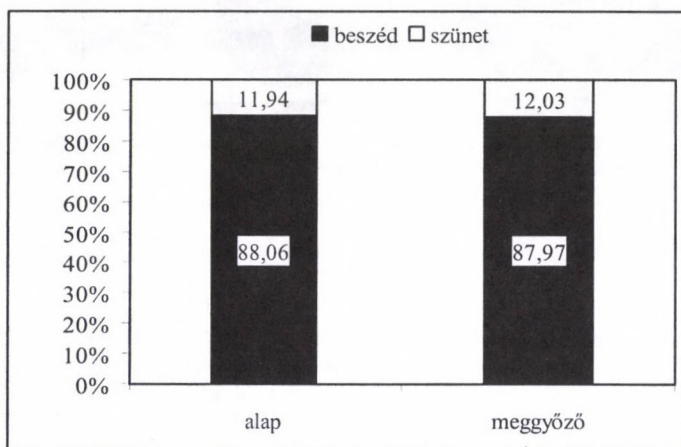
1. ábra
A beszédtempó változása

Az alapváltozatnál a kísérleti személyek átlagos beszédtempója 11,2 hang/s volt, a meggyőzőnél pedig 11,0 hang/s. Mint már az előbb utaltunk rá, ez a közel azonos átlagos beszédtempóérték nem jelenti azt, hogy az egyes beszélőknek ne változott volna a beszédtempója, egészen jelentős eltéréseket is találtunk a két felolvasás során megvalósuló beszédtempóértékek között. Az egyik női adatközlő beszédtempója 24%-kal csökkent, egy férfi kísérleti személyé pedig 16%-kal nőtt. Felmerül a kérdés, hogy hogyan realizálódtak ezek a változások; mely tényező változott, amely szerepet játszott a beszédtempó értékének létrejöttében. Az említett nő esetében a csökkenő beszédtempót a szünetek számának emelkedése okozta: az adatközlő az első változatban hétszer tartott szünetet, míg a másodikban már tizenhatszor. A beszédtempó növekedésének oka ellenben nem a szünetek számának változásában keresendő, hanem a szünetek időtartamának csökkenésében. Ugyanez a tendencia figyelhető meg a többi adatközlőnél is mind a gyorsítást, mind a lassítást tekintve. Ezek az adatok egybecsengenek Laczkó Mária (1993) eredményeivel, aki a tempó és a szünet viszonyát vizsgálta a hangos olvasásban. Kutatása alapján azt a következtetést vonta le, hogy „a tempó változtatásakor a gyorsítás főleg a szünetek időtartamának csökkenésében, a lassítás pedig a szünetek számának növekedésében jelentkezik” (188).

Feltételeztük, hogy jelentősebb különbség található az alap- és a meggyőző változat között a szünetekkel töltött idő mennyiségének tekintetében. Azt vártuk, hogy a felolvasás teljes időtartamán belül nagyobb lesz a szünetek aránya akkor, amikor a beszélők meggyőzők kívánnak lenni; hosszabb időt kitevő „hatásszüneteket” vártunk. Megmértük a felolvasások teljes időtartamát és az ezen belül szünettel töltött időtartamot, és kiszámítottuk a kettő arányát. A legtöbb adatközlőnél nem találtunk jelentős eltérést a két változat között, csak egyetlen kísérleti személynél volt az eltérés 10%-nál nagyobb. Az alap- és a meggyőző változatnál is átlagosan a teljes idő 12%-át tették ki a szünetek (2. ábra).

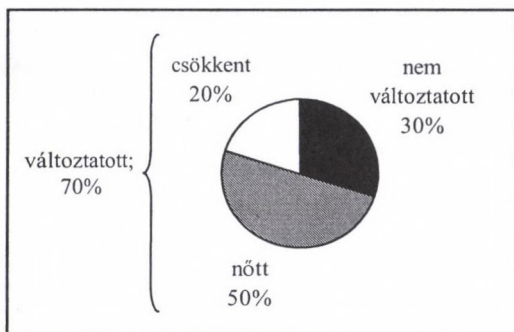
Nemcsak a szünetek időtartamában vártunk változást, hanem azok számában is, feltételeztük, hogy a meggyőző változatnál a beszélők több szünetet fognak tartani. Ez a hipotézisünk csak részben igazolódott; az adatközlők nagy részénél,

70%-ánál valóban változott a felolvasás során tartott szünetek száma, azonban csak a kísérleti személyek 50%-ánál találtunk több szünetet a meggyőző felolvasásnál, 20%-uknál csökkent a szünetek száma, 30%-uknál pedig nem tapasztaltunk e tekintetben eltérést (3. ábra).



2. ábra

A beszédjellel és a szünettel kitöltött idő aránya az alap- és a meggyőző változatokban

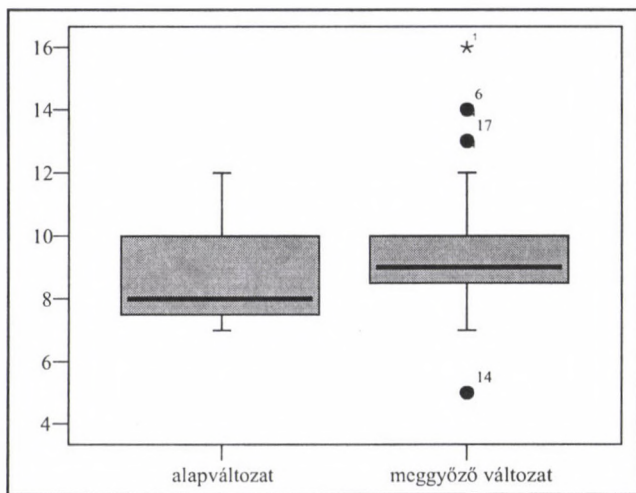


3. ábra

A szünetek számának változása

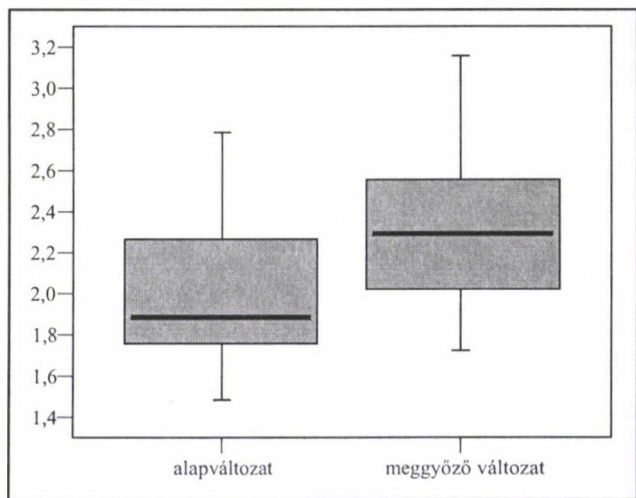
Az átlagos szünetszám az alapváltozatnál 9, a meggyőzőnél pedig 10 volt, tehát az eltérés statisztikailag nem releváns. Jelentős mértékben nőtt viszont a szórádás (lásd a 4. ábra „kilógó” értékeit). Ez azt jelenti, hogy az adatközlők a

meggyőzési szándékot a szünetek számának megváltozásával is kifejezték, de a változtatás iránya ezúttal sem tendenciaszerű.



4. ábra

A beszédsszünetek száma az alap- és a meggyőző változatokban



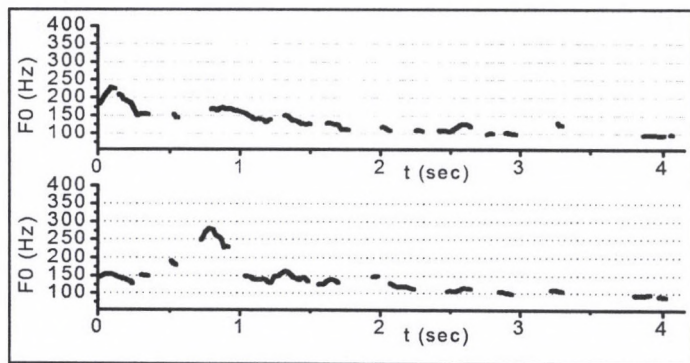
5. ábra

A hangterjedelem értékei az alap- és a meggyőző változatokban

Megmértük a kísérleti személyeknek a felolvasások során megvalósult hangterjedelmét, tehát azt a tartományt, amelyen belül az alaphangmagasságukat változtatták (vö. Gósy 2004: 187): az alaphangmagasság maximumát osztottuk az alaphangmagasság minimumával ($F_0\text{-max}/F_0\text{-min}$). Szignifikáns eltérés mutatkozott a két beszédváltozat között, a meggyőzés szándéka jelentősen megnövelte a beszélők hangterjedelmét ($t = 4,090$; $p = 0,0006$). Kiszámítható, hogy ezek az értékek milyen zenei hangközt jelentenek, így megkapjuk, hogy az alapváltozatnál az átlagos hangterjedelem alig egy oktáv, a meggyőzőnél pedig kb. egy oktáv és egy terc (5. ábra).

Az adatközlők többségére az volt a jellemző, hogy az alaphangmagasságuk minimuma kevésbé változott, jelentős változás az alaphangmagasság maximumának növekedésében jelentkezett – egyes beszélőknél közel 30%-os volt az eltérés a két változat alaphangmagasság-maximuma között. A kísérleti személyek többsége tehát felhasználta a meggyőzéshez hangja magasabb regisztereit is, magasabbra emelkedő dallámíveket is alkalmazott – ez egybevág a semleges, illetőleg emóciógazdag beszédmod között tapasztalható különbségekkel (vö. Vértés 1987).

E jelenség pontosabb leírásához megvizsgáltuk minden egyes mondat dallamszerkezetét. A beszédelemző szoftverrel kirajzoltattuk a mondat dallamgörbéit, majd minden adatközlőnél összehasonlítottuk ugyanazon mondatok két-két variánsát. Az alapváltozatokban a mondatok többsége ereszkedő dallammenetű volt, tehát a mondat elején volt a legmagasabb az alaphangmagasság értéke. A meggyőző változatoknál megnőtt azoknak a mondatoknak a száma, ahol a dallamszúcs a mondat belsejében volt található egy-egy fontosabbnak ítélt szó elején. A 6. ábrán ugyanannak a beszélőnek ugyanazt a mondatát láthatjuk, a felső diagramon az alapváltozatban, az alsón pedig a meggyőzőni szándékozóban.



6. ábra

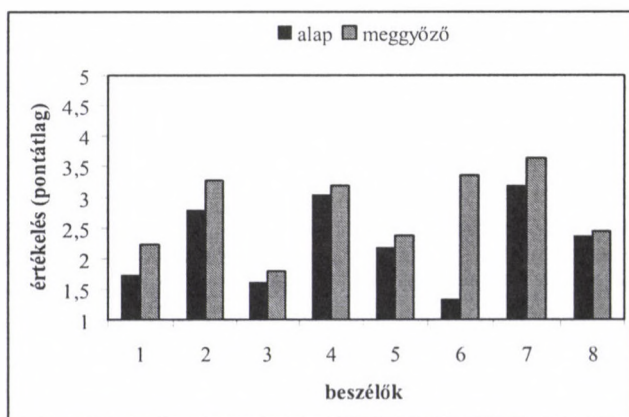
Ugyanazon beszélő által ejtett ugyanazon mondat az alap- (fent) és a meggyőző (lent) változatban („Mindössze három óra alatt odaérhet a varázslatos mediterrán városba.”)

Fontos megvizsgálnunk a hangsúlyozás különbségeit a két változatnál. A hangsúlyélményt több tényező is létrehozhatja, külön-külön vagy együttesen is, ezért a hangsúly jelensége nem mérhető pusztán egy akusztikai paraméterrel. A szakirodalom alapján létrejöhet *a)* a hangsúlyos szótag magánhangzójának intenzitását növelve, *b)* a hangsúlyos szótagban magasabb alaphangmagasság alkalmazásával, *c)* a hangsúlyos szótag magánhangzójának időtartam-növelésével, vagy *d)* a hangsúlyozandó szó előtti szünet tartásával (Gósy 2004: 198). Tehát a hangsúlyokat a legcélravezetőbb percepciósan vizsgálni; a hallgató mely helyeken érez hangsúlyt. Ebből a célból független megfigyelőt alkalmaztunk, akinek az volt a feladata, hogy a kinyomtatott szövegeken jelölje meg, hogy melyik szótagot hallja hangsúlyosnak. Az alap- és a meggyőző változatok között hangsúlyozás tekintetében szignifikáns eltérés mutatkozott ($t = 4,899$; $p < 0,0001$), a meggyőző változat során jóval többször éltek a hangsúllyal való kiemelés lehetőségével.

Percepciók kísérlet

Megvizsgáltuk, hogy a tesztelők egy-egy felolvasásnak milyen átlagos pontszámot adtak. Ebből az derült ki, hogy egyik változatot sem találták igazán meggyőzőnek, a legsikeresebb változatot is csak 3,63 átlagpontszámmal értékelték. Ez a férfi úgy változtatott beszédén a meggyőzés érdekében, hogy több szót emelt ki hangsúllyal, illetve ami a leginkább feltűnt a hallgatóknak: a hangterjedelme megnőtt, ezzel változatosabbá vált a felolvasása.

Ugyanakkor minden beszélő jobb eredményt ért el a meggyőzőnek szánt változatával, attól függetlenül, hogy a tesztelők az alap- vagy a meggyőző változatot hallották-e előbb (7. ábra). Nem minden kísérleti személy beszédváltozatai között egyformán nagy az eltérés, azok értek el nagyobb hatást, akik felolvasásukon a hangterjedelem vagy a szünetek időtartama tekintetében nagyobb mértékben változtattak.



7. ábra

A percepciók tesztelők értékelése beszélőnként

A tesztelők megjegyzéseiből kiderült, hogy nemcsak az említett szupraszegmentumok (hangterjedelem, hangsúlyok, hanglejtés, tempóviszonyok, szünetek) alapján értékelik az elhangzó felvételeket. A beszélők hangjára vonatkozó megjegyzésekkel találkozhattunk: *lágú, selymes, szép, kedves, mosolygós, bántó, kellemetlen, furcsa*. Ezek a kifejezések felvetik a fonetikai kísérletekben gyakran felmerülő kérdést: mi a jelentősége az egyéni hangszínezetnek. Mekkora szerepet játszik a meggyőzésben az orgánus? További kísérleteket igényel annak megállapítása, hogy a hallgatók mi alapján döntenek inkább: a hangszínezet vagy az egyéb szupraszegmentumok alapján.

Összegzés, következtetések

Az eredmények szerint a meggyőzés szándéka megváltoztatja a beszéd szupraszegmentális szerkezetét. Minden kísérleti személy változtatott a beszédén a meggyőzés érdekében, azonban ezek a változtatások nem mindegyik szupraszegmentum esetében azonos mértékűek és irányúak.

A beszéd időviszonyainak tekintetében nem tendenciaszerű a változtatás. A beszélők gyorsítanak vagy lassítanak: a lassítás a szünetek számának növekedésében, a gyorsítás a szünetek időtartamának csökkenésében mutatkozik meg. A kísérleti személyek kisebb része nem változtat beszéde időviszonyain.

Tendenciaszerű változtatást találtunk azonban a hangterjedelem, a beszéddallam és a hangsúlyozás vonatkozásában. A meggyőzésnél a beszélők hangterjedelme szignifikánsan nőtt, az F_0 maximuma megemelkedett. A beszéddallam vizsgálata alapján megállapíthatjuk, hogy az alapváltozatban a mondatok többsége ereszkedő dallammenetű, a meggyőző változatban viszont mondat belseji dallamsúcst tartalmaz. Az utóbbiakban számottevően több a figyelemfelkeltő hangsúly; a mondatok között kevesebb a kizárólag mondat eleji hangsúlyt tartalmazó, gyakoribb az egyes szavak kiemelése.

A percepciók tesztben az adatközlők egyik felvételt sem találták kiemelkedően meggyőzőnek. Ennek egyik lehetséges magyarázata az embereknek a reklámokkal szemben érzett tartózkodása. Érdekes eredmény, hogy már a 12-13 évesek is effajta szkepticizmussal viszonyulnak a direkt meggyőzéshez.

Azokat a változatokat találták a tesztelők a legmeggyőzőbbeknek, amelyekben jelentősen emelkedett a hangterjedelem az alapváltozathoz képest, illetve több kiemelő szerepű hangsúly jelent meg bennük.

Kísérleteink alapján az mondható, hogy a meggyőzésnek nincs egységes, mindenkre jellemző mintázata, a hallgatók mégis felismerik a meggyőzés szándékát. A beszélők valamilyen változtatással minden esetben elérik, hogy jobban hassanak a hallgatóságukra, amikor meggyőzőek kívánnak lenni. A beszélők meggyőzési stratégiái mindenképpen sikeresek lehetnek, még ha egészen különböző megoldási módokat választanak is.

Irodalom

- Adamik Tamás – A. Jászó Anna – Aczél Petra 2004. *Retorika*. Osiris Kiadó, Budapest.
- Arisztotelész 1999. *Rétorika*. (Fordította, a bevezetést és a jegyzeteket írta Adamik Tamás). Tólosz Kiadó, Budapest.
- Barthes, Roland 1997. A régi retorika. In Thomka Beáta (szerk.): *Az irodalom elméletei III*. Jelenkor Kiadó, Pécs, 69–168.
- Bóna Judit 2005. A hadaró és a gyors beszéd temporális sajátosságai. *Magyar Nyelvőr* 129. 235–242.
- Carnegie, Dale 1999. *Sikerkalauz 3*. Gladiátor Kiadó, Budapest.
- Cornificius 2001. *A szónoki mesterség. A C. Herenniusnak ajánlott retorika*. Magyar Könyvklub, Budapest.
- Cicero, Quintus Tullius 1987. *Válogatott művei*. Európa könyvkiadó, Budapest.
- Elekfi László – Wacha Imre é. n. [2003]. *Az értelmes beszéd hangzása*. Szemimpex Kiadó, Budapest.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Hull, Raymond 1997. *A sikeres nyilvános beszéd alapjai*. Bagolyvár, Budapest.
- Laczkó Mária 1993. A tempó és a szünet viszonya a hangos olvasásban. *Beszédkutatás* 1993. 185–193.
- Markó Alexandra 2005. A temporális szerkezet jellegzetességei eltérő kommunikációs helyzetekben. *Beszédkutatás* 2005. 63–77.
- Vértés O. András 1987. *Érzelm világunk és a nyelv történeti változásának kölcsönhatása*. Magyar Nyelvtudományi Társaság Kiadványai 179. Budapest.
- Wacha Imre 1994. *A korszerű retorika alapjai*. Szemimpex Kiadó, Budapest.

A SPONTÁN BESZÉD LEJEGYZÉSÉNEK NÉHÁNY MÓDSZERTANI KÉRDÉSE

Markó Alexandra – Bóna Judit

Bevezetés

A fonetikai vizsgálatok egyre gyakrabban tűzik ki célul a spontán beszéd vizsgálatát, aminek több oka is van. Egyik részről a tömegkommunikációs eszközök nagyobb mennyiségű és jobb technikai minőségű hanganyagot tesznek a vizsgálat számára hozzáférhetővé, mint néhány évtizeddel korábban; illetőleg a demokrácia nyitottabbá tett sok, korábban csak szűk csoportok számára elérhető nyelvhasználati színteret (például a hangzó médiát), ugyanakkor felszabadítólag hatott a beszédkedvre is (mind tartalmi, mind formai tekintetben). Másik részről a spontán beszéd vizsgálatához szükséges speciális eszközök (pl. többszatsornás felvételi lehetőségek, digitális elemző eljárások) elterjedése, hozzáférhetővé válása is megkönnyítette az ilyen jellegű kutatásokat. Ugyancsak fontos szempont, hogy a felolvasás, a művészi beszéd stb., azaz a nem spontán közlésformák elemzése egyre kevesebb nyitott kérdést hagy a kutatók számára. Ezek már inkább a másodlagos alkalmazások (elsősorban a beszédtechnológia) számára jelentenek fontos kutatási terepet, amelyek azonban a nyelvészeti-fonetikai mellett legalább annyira mérnöki szakértelmet is megkívánnak. Meg kell még említenünk a kutatói érdeklődést, illetőleg az objektivitásra törekvést is: a korábban csak személyes tapasztalatokból származó, sokszor kísérleti anyag vagy megfelelő módszer hiányában ellenőrizhetetlen feltételezéseket, megállapításokat immár a technikai feltételeknek köszönhetően kísérleti úton lehet igazolni vagy cáfolni.

A spontán beszéd lejegyzése a számítógépeknek a napjainkban tapasztalható mértékű elterjedéséig több nehézségbe ütközött, hiszen korábban sem a hangrögzítés, sem a lejegyzés technikai feltételei nem voltak kellően kidolgozottak. Ma a digitális hanganyagot és a lejegyzés eredményeként megjelenő írott szöveget erre a célra kidolgozott (és újabban már akár ingyenesen hozzáférhető) szoftveres háttérrel tudjuk egyszerre hallhatóvá és láthatóvá tenni – ugyanakkor természetesen az írott szöveg a hanganyagtól függetlenül is vizsgálható, kereshető és akár változtatható.

Az ELTE Fonetikai Tanszékén egy 2004 óta végzett kutatás keretében szisztematikusan rögzítettünk és dolgozunk fel spontánbeszéd-felvételeket, illetőleg a BME Távközlési és Médiainformatikai Tanszékével zajló együttműködésünk részét is képezik ilyen jellegű lejegyzési feladatok. Az itt olvasható módszertani elveket és gyakorlati problémákat így mintegy 100 órányi spontán beszéd lejegyzési tapasztalata alapján gyűjtöttük össze.

A tanszéki kutatás során a spontánbeszéd-lejegyzések célja egyrészt a spontán beszéd szupraszegmentális jellemzőinek a feltárása, másrészt – beszédtechnológiával foglalkozó szakemberekkel közösen – beszédfelismerő rendszerek betanítása számára jól használható transkripciók előállítására.

Az általunk használt két szoftver: *a)* az internetről ingyenesen letölthető Praat komplex akusztikai jelfeldolgozó (forrás: <http://www.fon.hum.uva.nl/praat/>; a legújabb verzió száma 4.5) és *b)* az ugyancsak ingyenes, kifejezetten szegmentálásra, címkézésre és lejegyzésre kifejlesztett Transcriber (a legújabb verzió száma 1.5; forrás: <http://trans.sourceforge.net/en/presentation.php>). Mindkét program felhasználóbarát grafikus felülettel rendelkezik, és többféle platformon (Windows, Unix) alkalmazható. A szoftverek angol nyelvűek, így a vezérlő felület, illetve a Transcriberben az automatikus címkék is angolul olvashatók. A lejegyzett szöveget a Praatban .txt, a Transcriberben .trs kiterjesztésű adatfájlokban tárolhatjuk és kezelhetjük. Az alábbiakban a lejegyzéshez szükséges előkészítő eljárásokat, majd a tanszéki munkálatok számára kidolgozott kétféle lejegyzési eljárást és kódrendszert mutatjuk be, illetve ezek előnyeit a további felhasználás számára.

A spontán beszéd szegmentálása és címkézése

A hanganyagok szegmentálása az első munkafázis – ennek folyamán bejelöljük a célnak megfelelően kiválasztott egységek határát. A címkézés vagy annotálás nem más, mint a szegmentált hanganyag szöveges tartalmának és/vagy a szükség szerinti (kiegészítő) információknak a rögzítése.

A tanszéki **szupraszegmentális kutatásokhoz** a szegmentálás a szöveges szakaszok és a szünetek, hallgatások (több résztvevő esetén a beszédjelnek két társalgási egység közötti kimaradása) határának kijelölését jelenti. A szöveges szakaszokra a *beszédszakasz* terminust alkalmazzuk, és két (néma vagy kitöltött) szünet (vagy egy szünet és egy hallgatás) közötti egységet értünk rajta. Beszédszakasznak tekintjük és jelöljük az önálló hümmögéseket is. A lejegyzés további céljától függően a szüneteken belül megkülönböztetjük a néma és a hangos szüneteket (az utóbbi esetben jelöljük a hezitálás jellegét: pl. *ő, öm*, köhintés, torokköszörülés stb.).

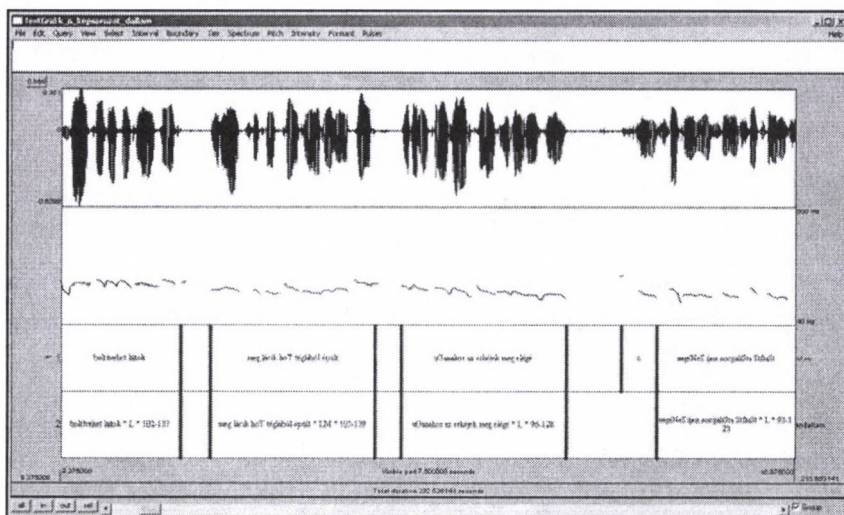
A szupraszegmentális vizsgálatok alapjául szolgáló spontánbeszéd-lejegyzésre alkalmasabbnak találtuk a Praat használatát, mivel többszintű kódolást is lehetővé tesz, és különösen mert a fonetikai vizsgálatokat is elvégezhetjük ugyanebben a szoftveres környezetben. A konkrét vizsgálati céltól függően tehát egy- vagy többszintű címkézési eljárást alkalmazunk. Az 1. ábrán egy monologikus szöveg annotációját láthatjuk két szinten (a Praat terminusával *tier*; szó szerint 'sor'): a szöveglejegyzést és a dallamcímkeket. A dialógusok esetében mindig legalább annyi szinten dolgozunk, ahány beszélő részt vesz a beszélgetésben/interjúban. Így a személyek beszédjellemei összevethetők.

A zöngétlen felpattanó zárhanggal és affrikátával kezdődő beszédszakaszok esetében a néma fázist, azaz a mássalhangzó artikulációjának részét képező, de

hallható akusztikai jellel nem járó időtartamot nem tekintjük a szünet vagy hallgatás részének. Ezekben az esetekben egységesen kb. 50 ms-nyi időt kivonunk a szünet időtartamából, illetőleg hozzáadunk a beszédszakaszéhoz, azaz a címke határát jelző jelölést ennyivel elmozdítjuk.

A nem magánhangzó-realizációként megjelenő [ə] hangot csak akkor címkézzük kitöltött szünetként, ha az [ə] képzését nem motiválhatta valamilyen fiziológiai kényszer (vö. Gósy 2004).

A címkézési szinteken a szándékolt vizsgálatától függően jelöljük a változókat. Az 1. ábrán látható példában a második szinten az alaphangmagasságra vonatkozó információkat jelöltük, mint a dallammenet és az F_0 minimum-, illetve maximumértéke. Itt jelöltük azt is, ha az illető beszédszakaszban volt nyikorgó zöngével (az angol szakirodalomban: *creaky voice*, a magyarban lásd még: *re-csegő zöngé, glottalizáció* stb.) realizált szakasz.



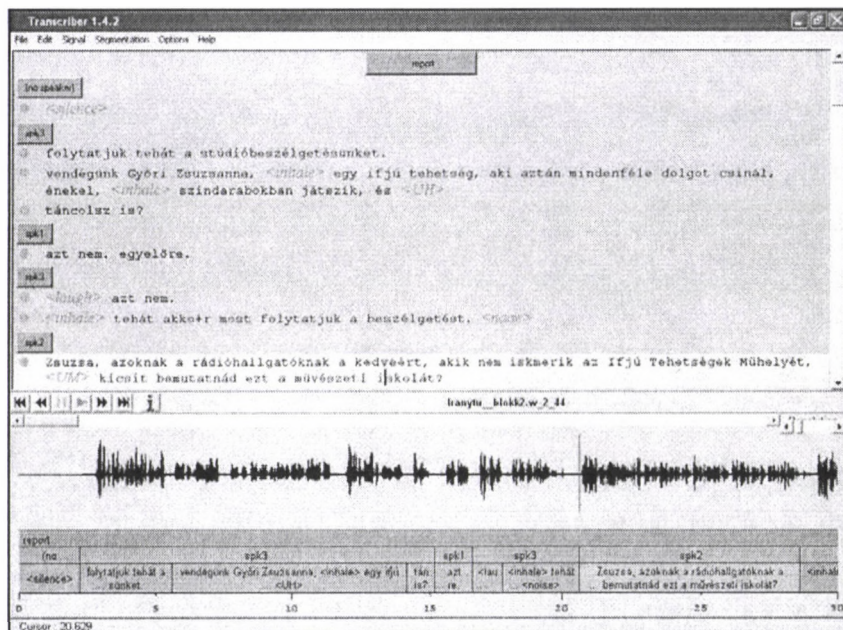
1. ábra

Az annotálás módja a Praat komplex akusztikai jelfeldolgozó szoftverben

A Transcribert a **beszédtechnológiai** felhasználás számára készülő lejegyzésekhez használjuk. Itt nagyobb egységeket szegmentálunk, általában a virtuális mondatok határait jelöljük be. A szüneteket a szegmentálás során nem mindig különítjük el a szöveges szakaszoktól, ezeknek nem is szükséges a pontos határát megjelölnünk – a beszédfelismerő miatt kis szünetet hagyhatunk a hangzó szakasz után. Amennyiben a mondat hosszabb, mint a címkézéskor három sornyi lejegyzés, akkor az egyik virtuális tagmondat határára tesszük a szegmentum-

jelet. Amennyiben a szakaszhatáron hallható a beszélő lélegzetvétele, ezt a következő szakasz elejéhez tartozónak jelöljük.

A program egyszintű címkézést tesz lehetővé, vagyis ugyanazon sorban szerepelnek a különböző beszélők megnyilatkozásai (2. ábra). A beszélőváltást a szövegjegyzést tartalmazó sor fölött láthatjuk (az első, második stb. beszélőre vonatkozóan: spk1; spk2 stb. – a *spk* a *speaker* szó rövidítése). Jelölhetjük viszont az egyszerre beszélést (overlapping speech), ilyenkor a lejegyzés is két sorban jelenik meg. Erre a beszédfelismerők szempontjából azonban nincs szükség, jelen munkálatunkban tehát érthetetlen <unintelligible> szakasznak címkézzük. (Itt és a továbbiakban csúcsos zárójelben az automatikus címkék olvashatók.)



2. ábra

Az annotálás módja a szegmentálásra és címkézésre kifejlesztett Transcriberben

A Transcriber nem alkalmas fonetikai mérésekre, viszont a beszédfelismerés szempontjából más beszédtevéyzők jelölésére igen. Meg kell különböztetnünk egymástól a néma szünetet <silence> (ennek külön jelölése csak akkor szükséges, ha a beszélő a rá jellemző egyéni szünettartásnál hosszabb, feltűnően hosszú szünetet tart); illetőleg azt, ha két beszédszakasz között belégzést <inhal>, kilégzést/lehelést <breath>, nyelvvel való csettintést <click>, szájzörejt <mouth>, köhögést <cough>, nevetést <laugh>, egyéb zajokat <noise> (a zaj

eleje és vége <noise_begin> és <noise_end>), valamint zenét <music> hallunk. A hezitálásokat és hümmögéseket is automatikus címkékkel jelöljük, az előbbiekre az <UH> és az <UM> jeleket, az utóbbiakra igenlés esetén az <UH-HUM> és az <UM-HUM>; tagadás esetén az <UH-UH> és az <UM-UM> jelek használatosak. Ezeknek a nem beszéd jellegű hangadásoknak a jelét a szövegjegyzésbe illesztjük bele. Az érthető suttogott beszédet a normál beszédhez hasonlóan jegyezzük le.

Lejegyzésmódok és jelölésrendszerek

A jegyzéshez a felhasználási céltól függően ugyancsak kétféle kódolást használunk. **A szupraszegmentális vizsgálatok alapjául szolgáló lejegyzés** számára saját „kvázifonetikus” rendszert dolgoztunk ki, amelynek kialakításában lényeges szempont volt, hogy különleges eszközt (szoftvert, betűkészletet stb.) ne igényeljen, valamint a lehető legnagyobb mértékben egyszerű és automatizálható legyen vele a hangok számának és így a tempóértékeknek a kiszámítása. Mindezek alapján sem a magyar helyesírás, sem a magyar nyelvű fonetikus lejegyzésre használatos két kódrendszer, az IPA és a magyar egyezményes jelölés egyike sem tűnt optimálisnak. A SAMPA (Speech Assessment Methods Phonetic Alphabet) számítógépes alkalmazásra megfelelő helyettesítője az IPA-nek (vö. <http://www.phon.ucl.ac.uk/home/sampa/>), és előnye, hogy létezik a magyar nyelvre kifejlesztett bővítése, azonban néhány jele a nemzetközi fonetikus lejegyzéshez szokott olvasó számára is nehézkes: pl. a 2 az [ø]-re, a d' a [j]-re vagy a J a [ɲ]-re (vö. <http://www.phon.ucl.ac.uk/home/sampa/hung-uni.htm>). A tempószámítás automatizálása céljából le kellett mondanunk a hosszúság külön karakterben történő jelöléséről is. Mivel azonban a spontán beszédben a hanghosszúság nehezen objektívizálható, és a hallás utáni lejegyzés a magasabb szintű percepciók működése miatt nem lehet pontos, ezt a szegmentális minőséget nem tartottuk fontosnak jelölni. Csak a magánhangzók esetében tüntetjük fel az időtartamot, mivel ott a magyar ábécé betűinek alkalmazása nem jár karaktertöbblettel. A betűkapcsolatokat is kiváltottuk, így a lejegyzett betűk száma megegyezik az általuk reprezentált hangok számával. Törekszünk a (relatív) könnyű olvashatóságra, így nem alkalmazunk számokat vagy más (nem betű) karaktereket. Mindezek miatt az egyértelmű betű-hang megfelelés nem teljes mértékű, de mivel a lejegyzett szövegek szegmentális fonetikai vizsgálata nem célunk, ennek nem is látjuk szükségét. A használt betűket az 1. táblázatban mutatjuk be és definiáljuk (a „Betű” oszlopokban az általunk használt betűjeleket, az „IPA-jel” oszlopokban a velük reprezentált hangzók nemzetközi fonetikus ábécébéli jelölését adjuk meg). A másodlagos képzési jegyekkel is jellemezhető beszédhangok közül csak azokat tüntetjük fel, amelyekre van példa a korpuszunkban. A mássalhangzó-hosszúságot nem jelöljük, így értelemszerűen a hosszú mássalhangzó-realizációkat nem soroljuk fel.

A lejegyzés a nemzetközi fonetikus jelölés elveit követi (vö. Handbook 1999). A szóhatárokat szóközzel jelöljük, de ahol a szóvégi és a szó eleji fonéma egyet-

len beszédhangban realizálódik, nem teszünk szóközt (például *máSodiképen* – értsd: 'második képen'). Központozást nem alkalmazunk.

1. táblázat: A lejegyzéshez használt jelek a szupraszegmentális elemzésre használt kódolásban

| Magánhangzók | | | |
|--------------|-----------|------|----------------|
| Betű | IPA-jel | Betű | IPA-jel |
| a | [ɔ] | o, ó | [o], [o:] |
| á | [a:] | ö; ő | [ø], [ə]; [ø:] |
| e | [ɛ] | u, ú | [u], [u:] |
| é | [e:] | ü, ű | [y], [y:] |
| i, í | [i], [i:] | | |

| Mássalhangzók | | | |
|---------------|--------------------|------|---------------|
| Betű | IPA-jel | Betű | IPA-jel |
| b | [b] | m | [m], [m] |
| c | [ts] | n | [n], [m], [ŋ] |
| C | [tʃ] | N | [ɲ] |
| d | [d] | p | [p] |
| D | [d͡z], [d͡ʒ] | r | [r] |
| f | [f] | S | [ʃ] |
| g | [g] | s | [s] |
| G | [ʒ] | t | [t] |
| h | [h], [ɦ], [ħ], [x] | T | [c] |
| j | [j], [ɟ], [ç], [ʝ] | v | [v] |
| k, K | [k], [kʰ] | z | [z] |
| l | [l] | Z | [ʒ] |

Alább egy ezzel a módszerrel lejegyzett szövegrészlet olvasható, a szakaszhatárokat és szünethelyeket | jelöli.

tehát akkor kezdem | az elSő képen eG | eGő tárSaSházát látok és | nem sámomra
elék furCa StíluSú mer az aja aján | boltíveket látok | meg láciK hoT téglából épült |
uGanakor az erkéjék meg elégē | ö | segēNeS ijen socijaliSta StíluSt | idéznek | tehát
eGenlőre est nem tudom | elhejezni Sem hejilek Sem időben | de a aNit látok hoG
eT forgalmaS hejen lehet | ez az épület | a máSodiképen | eT | embert látok ahoG
vaT fel v | Sőt eT kémēNSEprőt | ahoG vaT fel vaG | le | másik eG | falon | és még
láciK az ucán áló ember aki beszél vele és | úG úG nés ki mintha kioktatná | moSt
uGanez a | ö | kép vanaGon haSonló | de moSt a | fölfele másó kémēNSEprő | oktaTa
ki a | vaG legalábiS valmi | naGon | határozotó | tan álit valamit az ucán áló |
seméjnek aki | úG láciK hoG épen visakozik valami mijat | a neGediképen moS már
| Cak a | kémēNSEprőt látom ahoT fölfele másik | elég mérgeSnek tűnik | és már a |
úG láciK hoG úG gondolom hoG moS már magaSan járhat mer it látok eG erkéjt

melete | moSt a | kéméNSeprő | meglepetéSemre | bemásik eG | ő | erkéjre | ahol eG
 | söke nőt látok amint épen | út tünik hoT telefonál | és hátal ál neki tehát valószínűg
 eből valami meglepetés les | á igen és S | ésrevete a | férfit a kéméNSeprőt és moSt
 eGmáSt ölelik és naGom boldognak tünnek | és a | nő naGon kihívóan van öltözve

A számítógépes beszédfelismeréshez való felhasználásra készülő lejegyzésekhez alkalmazott kódrendszernek az alapja a magyar helyesírás, a szóalakok jelölése e szerint történik. Központozást is alkalmazunk, de a nagybetűs írásmódot csak a tulajdonnevek esetében tartjuk meg. Problémát a több szóból álló tulajdonnevek, például a műcímek (pl. *Az ember tragédiája*), a személynevek (pl. *Kossuth Lajos*), a több szóból álló intézménynevek (pl. *Fráter György Katolikus Gimnázium*) okozhatnak, de a jelen projekt során ezeknek a jelölésével nem foglalkoztunk külön.

A helyesírás nem tükrözi a kiejtésben megvalósuló koartikulációs szabályokat, de mivel ezek jó része a köznyelvi ejtésben általában érvényesül (pl. zöngességi hasonulás, összeolvadás), ezért a helyesírási alakok alapján a felismerésük automatizálható. A lejegyzés során olyan kiejtésvariációkkal is találkoztunk, amikor ezek a szabályok az elvárhatóságuk ellenére nem érvényesültek. Ezek bemutatására most nem térünk ki.

Azokban az esetekben, ahol a szóvégi és a szó eleji fonéma egyetlen beszédhangban realizálódik, a hangzó alakot normál helyesírással, két szóban jegyezzük le. Azt sem jelöljük külön, ha a szóvégi és a szó eleji azonos fonéma a beszélő ejtésében jól hallhatóan külön beszédhangban realizálódik.

A spontán beszéd sokszínűsége, ezen belül elsősorban a koartikulációs és a lazítási folyamatok működése miatt a nyelvi-fonológiai szerkezetek beszédbeli megvalósulásai eltérnek az írásban kódolt fonémasoroktól. A különféle kiejtési változatok a beszédfelismeréshez csak oly módon használhatók fel, ha a betanításhoz alkalmazott adatbázis jelölésrendszere tükrözi a nyelvi-fonológiai formát (is). Ezért a verbális jelek lejegyzésére a „fonetikus alak = nyelvi alak” megfeleltetést célszerű alkalmazni. Az egyértelmű megfeleltetés szempontjából lényeges, hogy az egybe tartozó (a kiejtésben összevont vagy egy szóhangsúllyal realizálódó) alakok jelölésében ne szerepeljen szóköz. A nyelvi egyértelműség ugyanakkor megkívánja, hogy az „egyenletek” jobb oldalán az írott nyelvi alak szabályos szóközeit jelöljük – erre a lejegyzésben az alsóvonalás jelet (_) használjuk: pl. *mer=mert, vót=volt, asszem=azt hiszem, nemtom=nem tudom*.

A spontán beszédben gyakori jelenség a szóvégi mássalhangzók (néhol szó eleji hangok) eltűnése, illetve a *-ban/-ben* és *-ba/-be* toldalékok felcserélése. Ezeket a hiányzó hangokat egy + jel segítségével kapcsoljuk a kiejtett szóalakhoz, például: *házba+n, azér+t*. Amennyiben egy hangnál több hiányzik a szóból, de egyértelmű, hogy minek kellene ott szerepelnie, kötőjellel, majd az egyenlőségjel után leírt helyes szóalakkal jegyezzük le. Például: *keresk=kereskedelmi*. Előfordul, hogy a beszélő szünetet tart szó közben, esetleg köhög, vagy zaj hallatszik. Ilyenkor a jelölés a következőképpen alakul: *lesz-*

=*lesznek* <*silence*> -*nek*. Problémás lehet a lejegyző számára, ha a névelő esetében fordul elő szünettartás, és a -z a következő szó elejéhez tapad. Például: *a=az* <*silence*> -z *infóláncban*; kevésbé jó megoldás, de esetleg elfogadható: *a=az* <*silence*> **-zinfóláncban**=*infóláncban*.

A szünettől és a hezitálástól eltérő megakadásjelenségeket különféleképpen jelöljük. Az egyszerű nyelvbottlást, metatézist tartalmazó szavakat megcsillagozzuk, és egyenlőségjellel összekapcsolva a hibás alak mellé leírjuk a helyes szóalakot is. Például: **kriállitás**=*kiállitás*. Hasonlóan csillaggal jelöljük a feltűnő hangzónyújtásokat, de ilyenkor nem szükséges a helyes alak megadása, a csillag és a betűkettőzés jelzi a hangidőtartam megváltozását. Például: *csalá*dd*oké*. Az újraindítást kötőjellel jelöljük, pl. *ho- holnap*; a téves kezdeket, ha nem tudni, milyen szót kezdtek volna, szintén kötőjellel jelölünk: *pi-*.

Az idegen szavak lejegyzésénél a fonetikus alakot és – ha ismert – a helyesírási szóalakot is lejegyezzük. A fonetikus alak után egyenlőségjelet teszünk, majd szögletes zárójelbe írjuk a helyesírási formát. Ha toldalék is kapcsolódik hozzá, azt / jellel jelöljük. Például: *vörksopra*=[*workshop*]/*ra*. Ha a lejegyző számára nem ismert az idegen helyesírású szóalak, akkor a fonetikus alakot teszszük szögletes zárójelbe, pl. [*mántenbájk*]. Hasonló a helyzet a hagyományos írásmódú személynevekkel, pl. *szécsényi*=[*Széchenyi*]. A spontán beszédben előfordulhat, hogy egy összetett szó előtagja idegen szó, utótagja magyar szó. Ilyenkor is a helyesírás szabályai a mérvadóak, például *lazakne*=[*lasagne*]*tésztát*; de elfogadható a *lazaknetésztát*=[*lasagne*]*tésztát* lejegyzés is.

A két jelölésrendszer fő különbségei jól megfigyelhetők, ha összevetjük ugyanannak a szövegnek a szupraszegmentális elemzéshez és a beszédfelismerő betanításához készült leiratát. Ez utóbbi olvasható a következőkben:

tehát akkor kezdem. <inhal> az első képen egy <inhal> egy társasházat látok, és <inhal> ne*mm*, számomra elég furcsa stílusú, mer=mert az alja alján <inhal> boltíveket látok, <inhal> meg látszik, hogy téglából épült, <inhal> ugyanakkor az erkélyek meg eléggé <silence> <click> <silence> <UH> szegényes ilyen szocialista stílust <noise> idéznek, <inhal> tehát egyelőre ezt nem tudom <breath> elhelyezni sem helyileg, sem időben, <silence> <click> <inhal> de a- annyit látok, hogy egy forgalmas helyen lehet <silence> ez az épület. <silence> <noise> <silence> <inhal> <silence> a második képen <inhal> egy <noise> embert látok, ahogy vagy fel, v- <noise> <silence> sőt egy kéményseprőt, <inhal> ahogy vagy fel, vagy <noise> le <noise> mászik egy <noise> falon. <inhal> és még látszik az utcán álló ember, aki beszél vele, és <silence> úgy úgy néz ki, mintha kioktatná. <noise> <silence> <noise> <silence> <noise> <silence> most ugyanez a <silence> <UH> <breath> <silence> kép van, nagyon hasonló, <inhal> de most a <inhal> fölfele mászó kéményseprő <inhal> <silence> oktatja ki a <noise> vagy legalábbis valami <noise> nagyon <noise> határozott- <breath> -tan állít valamit az utcán álló <inhal> személynek, aki <silence> úgy látszik, hogy éppen visszakozik valami miatt. <noise> <silence> a negyedik képen mos=most már <inhal> csak *aa* <breath> <silence> kéményseprőt látom, ahogy fölfele mászik, <inhal> elég mérgesnek tűnik, <inhal> <silence> és már a <silence> úgy látszik, hogy úgy

gondolom, hogy mos=most már magasan járhat, mer=mert itt látok egy erkélyt mellette. <silence> most a <breath> kéményseprő <silence> meglepetésemre <noise> <silence> bemászik egy <silence> <noise> <silence> <UH> erkélyre, <inhale> ahol egy <silence> szőke nőt látok, amint éppen <inhale> <silence> úgy tűnik, hogy telefonál, <silence> <noise> <silence> és háttal áll neki, tehát valószínűleg=valószínűleg ebből valami meglepetés lesz. <silence> á, igen, és <silence> észrevette a <breath> <silence> férfit, a kéményseprőt, és most egymást ölelik, és nagyon boldognak tűnnek, <inhale> <silence> és a <breath> <silence> <noise> <silence> nő nagyon kihívóan van öltözve.

Az összehasonlítás kedvéért ugyanezt a szöveget bemutatjuk helyesírással jelölve (egy lehetséges központoszással) is:

Tehát akkor kezdem. Az első képen egy egy társasházat látok, és nem, számomra elég furcsa stílusú, mert az alja alján boltíveket látok, meg látszik, hogy téglából épült, ugyanakkor az erkélyek meg eléggé ő szegényes ilyen szocialista stílust idéznek. Tehát egyenlőre ezt nem tudom elhelyezni sem helyileg, sem időben, de annyit látok, hogy egy forgalmas helyen lehet ez az épület. A második képen egy embert látok, ahogy vagy fel, v- sőt egy kéményseprőt, ahogy vagy fel-, vagy lemászik egy falon, és még látszik az utcán álló ember, aki beszél vele, és úgy úgy néz ki, mintha kioktatná. Most ugyanez a ő kép van, nagyon hasonló, de most a fölfele mászó kéményseprő oktatja ki a vagy legalábbis valami nagyon határozottan állít valamit az utcán álló személynek, aki úgy látszik, hogy éppen visszakozik valami miatt. A negyedik képen most már csak a kéményseprőt látom, ahogy fölfele mászik, elég mérgesnek tűnik, és már a úgy látszik, hogy, úgy gondolom, hogy most már magasan járhat, mert itt látok egy erkélyt mellette. Most a kéményseprő meglepetésemre bemászik egy ő erkélyre, ahol egy szőke nőt látok, amint éppen úgy tűnik, hogy telefonál, és háttal áll neki, tehát valószínűleg ebből valami meglepetés lesz. Á, igen, és észrevette a férfit, a kéményseprőt, és most egymást ölelik, és nagyon boldognak tűnnek, és a nő nagyon kihívóan van öltözve.

Összegzés

A két jelölésrendszer legfontosabb eltérései: míg a szupraszegmentális elemzésekhez készülő lejegyzés kvázifonetikus (általános szabályai a nemzetközi fonetikus lejegyzés elveinek felelnek meg), csak a kiejtésre összpontosít, és nem vonatkoztatjuk a magyar helyesírás szerinti írásképre; addig a beszédfelismerő betanításához készülő jelölés a magyar helyesírásból indul ki, és a lehető legnagyobb mértékben tükrözi a kiejtés és a helyesírás szerinti jelölésmód eltéréseit. Míg az előbb bemutatott lejegyzésben csak a szupraszegmentális vizsgálat céljait és elveit vesszük figyelembe, ezért kiemelt fontosságú a szegmentumok száma (a tempószámítások megkönnyítésére), a szünetek helyének, időtartamának és néma vagy kitöltött voltának jelölése, a dallamsajátosságok kódolása stb.; addig az utóbbi céljának megfelelően a fonetikai szegmentumszám, a szünet pontos hossza stb. mellékes, ugyanakkor a szöveges részek közötti egyéb információk

(ki- és belégzés, nevetés stb.) jelentőséggel bírnak, ezért jelölendők. A kétféle jelölés céljainak más-más szoftveres támogatás felel meg jobban: a fonetikai elemzésre is alkalmas, többszintű címkézést lehetővé tevő Praatot alkalmazzuk a hangtani alap kutatásban; a Transcriber azonban a beszédtechnológiai alkalmazás számára ad megfelelőbb eszközháttérrel.

Irodalom

Gósy, Mária 2004. The manifold function of schwa. *Grazer Linguistische Studien* 62. (Herbst 2004) 15–26.

Handbook 1999 = *Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet*. Cambridge University Press, Cambridge.

A szerzők köszönetet mondanak Fegyő Tibornak és Mihajlik Péternek értékes tanácsaikért, illetve az AITIA International Zrt.-nek és a BME TMIT-nek, amelyek lehetővé tették a beszédtechnológiai projektekből való részvételüket.

A tanulmány az alábbi pályázatok támogatásával, illetve projektek keretében jött létre:
OTKA T046455 számú pályázat;

MALACH (Multilingual Access to Large Spoken Archives from the Holocaust; támogatók: Survivors of the SHOAH Visual History Foundation and the Johns Hopkins University in Baltimore [Agreement JHU 8202–48279], The National Science Foundation [Grant No. IIS–0122466])

és

Voxearch – digitális médiaarchívumok automatizált katalogizálása beszédfelismerés segítségével

(támogatók: GVOP–3.1.1–2004/05–0385/3.0 pályázat, Szent István Rádió).

A SPONTÁN BESZÉD ÉS A BESZÉDFELDOLGOZÁS ÖSSZEFÜGGÉSEI GYEREKEKNÉL

Horváth Viktória

Bevezetés

Az anyanyelv-elsajátítás során a beszédprodukción és a beszédpercepción párhuzamosan fejlődik. Az első szakaszban a csecsemő minta nélkül kezd gögicsélni, spontán érés eredményeképpen. Az akusztikai visszacsatolás, majd a környezetből jövő emberi hang újra és újra kiváltja, ösztönzi a hangadást. Ezt követően a nyelvelsajátítás utánzással valósul meg (Beke 1996). Az utánzás sikerességében nagyon fontos szerepe van a pontos észlelésnek: a gyermek által létrehozott hangsor csak ennek megfelelő működése esetén hasonlít a felnőtt nyelvi mintához. A megfelelő feldolgozás tehát elengedhetetlen feltétele a beszéd megindulásának. Ép fejlődés esetén a percepción minden életkorban magasabb szintű, mint a produkcion. Az első években a beszédprodukción fejlődése gyorsabb és látványosabb; két-három éves kor után azonban a fejlődés mértéke kiegyenlítődik (Gósy 2005). A 6-7 éves gyermekek spontán beszédében a toldalékolás már a felnőtt nyelvi mintának megfelelően történik. Fokozatosan eltűnik az addigi életkorra jellemző téves analógia. Hatéves kor után gyakoribbá válnak az összetett mondatok, a gyermekek egyre bonyolultabb szerkezetekkel fejezik ki magukat. Többségük már hosszabb, monologikus közlések létrehozására is képes; jelentősen növekszik beszédükben a használt szavak és kifejezések száma. A percepción fejlődését tekintve a vizsgált életkorban – ép fejlődés esetén – a gyermekek képesek megkülönböztetni a beszédhangokat, bár az időtartam felismerése még nem minden esetben hibátlan. A beszéd szupraszegmentális tényezőinek feldolgozása még ebben az életkorban is fejlődik.

Különböző tényezők azonban zavart eredményezhetnek a fejlődésben, ami érintheti a produkciót, a percepciót, esetleg mindkét folyamatot. A beszédrendszer organikus zavara, az idegrendszer zavarai, a verbálisan ingerszegény vagy a kódváltó környezet, a pszichés problémák eredményeként létrejött beszédzavarok megjelenési formái igen változatosak lehetnek (Gósy 2000). A beszédképzés hibái feltűnőek, ezért szakember segítségével a gyermek problémái időben rendezhetők. A percepción zavarai azonban akár évekig is rejtve maradhatnak – a gyermek sokféle kompenzációs stratégiát használ az elhangzottak feldolgozásához –, így sok esetben csak akkor diagnosztizálják a problémát, ha már nehéz rajta segíteni. Megfelelő terápia hiányában ezek a gyermekek sorozatos kudarcokat élnek át az iskolában, amelyek később magatartászavarokhoz és deviáns viselkedéshez vezethetnek (Schneider–Simon 1996).

A feldolgozás problémái felnőttkorban is megnehezítik a mindennapi kommunikációt, az ismeretszerzést, akár a munkavégzést is. Ezért nagyon fontos a perceptiós zavarok mielőbbi diagnosztizálása és a fejlesztés megkezdése.

Pedagógiai tapasztalat, hogy ha a gyermek nagyon keveset beszél, s közlései leginkább csak néhány szóból álló szerkezetek, akkor a mentális lexikon bővülésének lehetőségei korlátozottak, a tanulás folyamata valószínűleg nehezített. Ezzel szemben, ha a gyermek szívesen beszél, a felnőttekhez hasonlóan sok és sokféle összetett közléseket produkál, fel sem merül annak a gyanúja, hogy a produkció megfelelő fejlődése mellett a gyermek esetleg perceptiós elmaradással küzd. Mindezek felvetik a kérdést, hogy vajon mi az összefüggés a két folyamat között. Az életkornak megfelelő szintű beszédprodukció jelezheti, hogy a perceptiós folyamatok is jól működnek; de az is előfordulhat, hogy a jól működő beszédképzés ellenére gondok vannak a feldolgozásban. A produkcióban jelentkező elmaradás sok esetben éppen a perceptiós zavarra vezethető vissza, hiszen ha a gyermek nem észleli pontosan az elhangzottakat, akkor nem is fogja tudni azt helyesen reprodukálni. Természetesen az is előfordulhat, hogy a beszédhibák ellenére a feldolgozás az életkornak megfelelően működik.

Tekintettel a beszédfeldolgozás közvetlenül nem tapasztalható folyamataira, felmerül a kérdés, hogy gyermekkorban a spontán beszéd mutat-e valamit a feldolgozás mikéntjéről: találhatunk-e a gyermek beszédében olyan jelzéseket, amelyek utalnak a fennálló perceptiós elmaradásra. A jelen kutatás – magyar nyelven elsőként – arra a kérdésre keresi a választ, hogy vajon a spontán beszéd folyamatosságának és komplexitásának mértéke jelezheti-e a gyermek észlelési és megértési működéseinek zavarait; illetve a perceptiós elmaradás együtt jár-e a tervezési fejlettség alacsonyabb szintjével.

A kiinduló hipotézis szerint a gyermekek produkciójának és perceptiójának szintje összefüggést mutat. Azt feltételeztük, hogy minél összetettebb közléseket produkál a gyermek, annál valószínűbb, hogy a perceptiós folyamatai jól működnek. A feldolgozás sikeressége ugyanis feltétele a mentális lexikon bővülésének, a bonyolultabb grammatikai szerkezetek elsajátításának. Vagyis ha a gyermek szinte kizárólag főneves szerkezeteket vagy egyszerű mondatokat produkál a vizsgált életkorban, kevés bővítményt használ, akkor gyanítható, hogy perceptiós zavarok húzódnak a háttérben. Feltételeztük, hogy a beszéd szerkezeti komplexitása és a megakadásjelenségek száma összefügg: a bonyolultabb szintaktikai szerkezetek várhatóan több hibázással járnak együtt. Ez arra vezethető vissza, hogy a közlés hosszú és összetettsége fontos tényezői annak, hogy a beszéd folyamatosságát megszakítja-e valamilyen megakadásjelenség (McLaughlin–Cullinan 1989, Yaruss–Newman–Flora 1999).

Kísérleti személyek, anyag és módszer

A kísérletben 9 fiú és 7 lány vett részt; mindannyian 6-7 évesek, ép hallók, és nem beszédhibások. A gyermekek hasonló szociális környezetben nőttek fel, és időben kezdtek beszélni. Az óvodásokkal interjúk készültek arról, hogy hol töl-

tötték a nyarat, mivel szoktak játszani, vagy melyik a kedvenc meséjük a tévében. A közlések rögzítése a megszokott óvodai környezetben, magnetofonnal történt. A teljes felvétel időtartama 46 perc. Természetesen nagy különbségek voltak a gyerekek között abban, hogy hány percet és milyen folyamatossággal beszéltek. A legrövidebb interjú időtartama 1'00", a leghosszabbé 6'20". Az óvodások közül sokan hosszú, összetett közléseket produkáltak, de voltak olyanok is, akik nyári élményeiket egy-egy szóval, felsorolásszerűen adták elő, például: *Játszottam. Fürödtem. Aludtam.* Ilyenkor a kísérletvezető segítő kérdésekkel próbálta bővebb közlésre ösztönözni a gyermekeket.

A spontán beszéd vizsgálata egyrészt arra irányult, hogy milyen a gyermekek beszédének szintaktikai komplexitása. Ennek elemzéséhez az ún. KFM-módszert használtuk (Gerebenné-Gósy-Laczkó 1992). Az amerikai Lee és Canter a hetvenes években kidolgozott egy olyan kritériumrendszert, amely alapján a gyermek spontán beszédének komplexitása megítélhető. Az eljárás azt vizsgálja, hogy milyen a gyermek beszédének morfológiai és szintaktikai komplexitása; a mondatok struktúrája és hosszúsága; a névmások és egyéb szófajok, valamint az igeidők, a kérdő és a tagadó formák használatának szintje. (A módszer elméleti hátterét Chomsky generatív grammatikája adta.) A vizsgálatot elsősorban 3–7 éves gyermekek számára dolgozták ki. A fő cél az volt, hogy klinikai vizsgálati helyzetekkel szemben természetes körülmények között legyen megítélhető a gyermek beszédteljesítménye (a szorongás ugyanis nagymértékben csökkentette az óvodások teljesítményét), vagyis amikor a gyermek kötetlen témáról beszélget a felnőttel – csakúgy, mint a hétköznapiakban.

A magyar nyelvre Gerebenné Várbíró Katalin, Gósy Mária és Laczkó Mária adaptálta a tesztelések eljárást, amelyet a **közlesegységek fejlődési mutatójaként** (KFM) neveztek meg. Az értékelés lényege a következő: a gyermek rögzített spontán beszédének lejegyzését követi annak megállapítása, hogy a beszédben milyen pontértékű szavak és szerkezetek vannak. Ez azt jelenti, hogy a spontán beszéd szavainak, szerkezeteinek és mondatainak meghatározott pontértéke van. A szavak pontozásának alapelve a magyar nyelvtan sajátosságait, illetve a magyar gyermek anyanyelv-elsajátítási szakaszait veszi figyelembe. Nem minden szófaj kap pontot, csak a névmások, számnevek, határozószók, névutók és kötőszók. (Ha például a gyermeknek felteszik a *Ki ment el?* kérdést, és a válasza: *Anyu*, erre nem kap pontot. Ha azt mondja: *Ő*, erre 1 pont jár.) A kritériumrendszer azt is meghatározza, hogy az egyes szófajokon belül bizonyos csoportok milyen pontértékűk van.

A személyes névmás használatára például 1 pont jár, a visszahatóra 3, a kölcsönösre 4. A szerkezeteken belül külön pont jár a ragozásra, és külön a bővítményekre is. A *filmet néztem* szerkezet ennek megfelelően a múlt idejű tárgyas ragozás miatt 4 pontot ér, és még 1-et az állítmány és a tárgy használata miatt. A ragozás rendszerén belül a legegyszerűbb a jelen idejű alanyi ragozás (1 pont), a múlt idejű tárgyas ragozás 4 pontot ér, az ige és a főnévi igenév használata 5 pontot (*lehet enni*), a gyermek a legtöbb pontot a múlt idejű feltételes mód

használatára kapja (*adtam volna* – 9 pont). A bővítmények pontozása is az egyszerűbbtől a bonyolultabb felé halad: állítmány és tárgy (1 pont), állítmány és hely- vagy időhatározó (3 pont), állítmány és állandó határozó (*beszélt a mamáról* – 6 pont). A tagadó formák is pontozva vannak komplexitásuk szerint 1–3 pontig. A kötőszók 6 kategória szerint vannak csoportosítva, a kapcsolatos 1 pontot, a feltételes 6-ot ér. A mellékmondatban szereplő felszólító forma használatát is külön pontot ér; amelyben tiltó forma van, magasabb értékkel jár (*azt mondtam, hogy ne kiabáljatok* – 4 pont).

A szavak és a szerkezetek értékelése után az egyes mondatok is pontot kapnak. (A spontán beszéd mondatokra tagolása – mint jól ismert – több problémát is felvet, bár a gyermekek rövidebb közléseiben viszonylag egyszerűbb volt a mondathatárok szubjektív megállapítása.) Minden nyelvtanilag helyes mondat 1 pontot ér, vagyis amely nem tartalmaz grammatikai hibát (szám és személy egyeztetése, névelő és szókezdő hang összehangolása stb.). Fontos elv, hogy a nyelvtanilag hibás mondatok szavai és szerkezetei kapnak pontot, csak maga a mondat nem. A KFM-érték kiszámítása tehát a következőképpen történik: a szavak és a szerkezetek értékeléséből összejött pontszámot hozzá kell adni a helyes mondatok számához, majd ezt a teljes összeget elosztani az összes mondat számával. A jelen kutatásban ezzel a módszerrel elemeztük minden óvodás beszédének grammatikai komplexitását. A teszt sztenderdizálása magyar nyelvre még nem történt meg. Az adaptálás során a magyar kutatók azonban minden korosztályra megadtak egy tapasztalati értéket, amely az adott életkorban elvárható. Ez a KFM-érték 6 éves korban 11.

A spontán beszéd vizsgálatának másik aspektusa a megakadásjelenségek elemzése. A beszéd folyamatosságát megszakíthatják olyan jelenségek, amelyek a tervezés bizonytalanságának következményei, ilyen például a hezitálás vagy a töltelékszavak használata. A másik nagy csoport a téves kivitelezések: ezek a tervezés egy vagy több szintjén megjelenő hibák. Osztályoztuk és összesítettük a korpuszban található bizonytalanságból adódó megakadásokat és a hibákat. A felvételek eltérő hosszúra való tekintettel a percnként előforduló megakadások számát hasonlítottuk össze, és azt vizsgáltuk, hogy az óvodások átlagosan hány megakadásjelenséget produkáltak percnként. Az egyes jelenségek típusba sorolásakor a Gósy Mária által kidolgozott legújabb kategorizációt vettük alapul (2004). Elemeztük, hogy ebben az életkorban a gyermekek spontán beszédében mely megakadástípusok fordulnak elő leggyakrabban; illetve melyek a nyelvi tervezés azon szintjei, amelyekkel a gyermekeknek a legtöbb nehézségük van.

A beszédpercepciót a sztenderdizált, magyar fejlesztésű GMP-diagnosztika (Gósy 2006) segítségével mértük fel. A diagnosztika a beszédfeldolgozási rész-folyamatok működését vizsgálja a hallástól az értelmezésig. A tesztekkel felmérhető a zavar típusa is. A jelen kutatásban hét részteszt eredményeivel dolgoztunk, amelyek az akusztikai, fonetikai és fonológiai észlelést, a szerialitást, a mondat- és a szövegértés képességét vizsgálják. Az akusztikai észlelés tesztelése zajjal fedett mondatok (GMP2), illetve szavak (GMP3) segítségével történik: a

gyermeknek a magnetofonról hallottakat kell megismételnie. A fonetikai feldolgozás – amely a hallás és az észlelés megfelelő összefüggése – vizsgálata (GMP4) során a gyermek szűk frekvenciájú (2200 és 2700 Hz közötti) mondatokat hall a magnetofonról, amelyeket meg kell ismételnie. A fonológiai észlelés működésének vizsgálata (GMP5) gyorsított (15 hang/s) mondatok visszamondatásával történik. Itt a mondatok szerkezete és jelentéstartalma meghaladja ennek a korosztálynak a nyelvi képességeit, mert nem a mondatok megértésének a tesztelése a cél, hanem az észlelési teljesítmény megértéstől független ellenőrzése. A szerialitás felismerésének készségét értelmetlen hangsorok visszamondatásával vizsgálják (GMP10): a gyermek képes-e az elhangzott beszédhangok és sorrendjük helyes azonosítására. A szövegértési feladat egy mese meghallgatásából, majd 10 megértést ellenőrző kérdésből áll (GMP12).

A mondatértés tesztelése (GMP16) pedig a következőképpen történik: a gyermek két – egymástól csak egy részletben különböző – képet lát maga előtt, majd egy mondat elhangzása után rá kell mutatnia arra a képre, amelyre a mondat vonatkozik.

A gyermekeket délelőtt, a spontán beszéd felvétele után óvodai környezetben teszteltük. Az összehasonlíthatóság miatt az egyes feladatoknál kapott eredményeket átlagolva egy általános percepciósi értéket (PÉ) is megállapítottunk minden óvodásnál. Az adatok statisztikai elemzését (Pearson-féle korrelációelemzés) a 11.0 verziószámú SPSS szoftver segítségével végeztük.

Eredmények

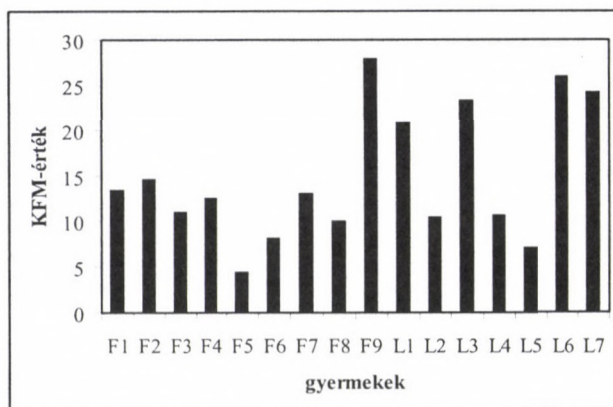
A spontán beszéd jellemzői: grammatikai komplexitás

A jelen kutatásban részt vett gyermekek meghatározott közlésegségei fejlődési mutatóinak átlagos pontszáma 14,84, ami azt jelenti, hogy az óvodások bonyolultabb jelzős, határozós szerkezetekkel, névmások és számnevek gyakori használatával, főnévi igeneves szerkezetekkel, jelen és múlt idejű alanyi és tárgyias ragozás megfelelő alkalmazásával fejezték ki magukat (egy kislány nyári élményeiből: *Az új csúszdán csúsztunk, habár olyan sokszor, hogy, hogy ott volt egy egész tömeg, h-, nem bírtunk sorakozni, aztán most kerget minket a Rudi...*). A kapott adatok nagy egyéni különbségeket mutatnak (1. ábra).

Az óvodások 62,5%-a jól beszélőnek számít az alkalmazott kritériumrendszer alapján, 6 gyermek viszont nem érte el az ebben az életkorban elvárható, tapasztalatilag megállapított 11 pontot. Pszicholingvisztikai kísérletek igazolták, hogy a lányok többségének anyanyelvi szintje az egyes életkorokban általában magasabb, mint a fiuké; ugyanakkor a fiúk teljesítenek rendszerint a legjobban és a leggyengébben (Gósy 2005). A különbség egyrészt hormonális okokkal, másrészt a női és a férfiagy morfológiai különbségeivel magyarázható (Papp 2006). Ezek alapján azt vártuk, hogy a lányok beszédét nagyobb mértékű grammatikai komplexitás jellemzi. Az eredmények igazolták a hipotézist: a lányok több mint felének beszédét 20 vagy annál magasabb KFM-érték jellemezte, ami azt mutat-

ja, hogy hosszan és szívesen beszéltek, bonyolultabb morfológiai és szintaktikai szerkezeteket alkalmaztak.

A fiúkra nem volt jellemző a nagyon magas KFM-pontszám, közléseik egyszerűbbek és bővítményekben szegényebbek voltak; gyakoribbak náluk a tömörebb, sok esetben felsorolásszerű válaszok. A fiúknál találtuk a két szélső értéket, KFM = 4,51 és 27,88 pont.



1. ábra

A KFM-értékek a fiúk (F) és lányok (L) spontán beszédében

A spontán beszéd jellemzői: megakadásjelenségek

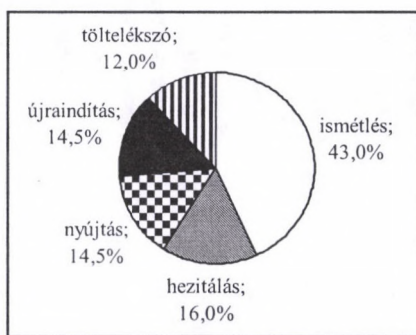
A 46 perces anyagban 388 megakadást találtunk; az elvártaknak megfelelően mind a 16 óvodás rögzített anyagában szerepelt valamilyen megakadásjelenség. A gyermekek átlagosan 7 megakadást produkáltak percenként – felnőtt adatközlőknél átlagosan 5 fordult elő egy perc alatt egy irányított spontán beszédet vizsgáló kísérlet során (Horváth 2004). A teljes korpuszban 18 típusra találtunk példát az elemzett huszonegyből; nem fordult elő malapropizmus, freudi elszólás és fonológiai hiba.

A megakadásjelenségek két nagy csoportját tekintve a gyermekek spontán beszédében található megakadások 35,8%-a sorolható a téves kivitelezések körébe; tehát a jelenségek nagyobb része a beszélők bizonytalanságából adódott (64,2%). A felnőttek narratíváiban (Gósy 2003) nagyon hasonló arányokat találunk (bizonytalanság: 67,24%, hiba: 32,76%).

Az eredmények azt mutatják, hogy az óvodások bizonytalansága főként ismétlésekben nyilvánult meg (43,0%). Ennél jóval kevesebb, és nagyon hasonló arányú a hezitálás (16,0%), a nyújtás (14,5%), az újraindítás (14,5%) és a töltelészó (12,0%) előfordulása (2. ábra).

Ha tehát a gyermek elbizonytalanodik a tervezési folyamat során, akkor leggyakrabban azt a stratégiát választja, hogy megismétli az addig elhangzottakat, s

ezzel nyer időt a továbbiak átgondolására. A felnőttek spontán beszédében a hezitálás a leggyakoribb a bizonytalanságok közül (Gósy 2003). A felnőttek közlései – és ennek megfelelően a tervezési részfolyamatok – már jóval bonyolultabbak, így nekik egyszerűbb stratégiára van szükségük az „időnyeréshez”. Az ismétlés tehát nem megfelelő stratégia, mert fonológiai, fonetikai tervezést és kivitelezést igényel, amely figyelmet von el a tervezés felsőbb szintjeitől. Ezzel szemben a hezitálás rendszerint csupán egy hang rövidebb-hosszabb ejtésével realizálódik (nem történik tervezés, csak artikulációs kivitelezés), így közben a tervezési folyamatokra nagyobb figyelem fordítható. Feltételezhető, hogy amint a gyermek közlései egyre összetettebbek lesznek, az ismétlés már nem elég hatékony a tervezési bizonytalanságok feloldására, így a felnőtt mintát követve a gyermek is inkább a hezitálást fogja előnyben részesíteni.



2. ábra

Az egyes típusok előfordulási aránya
a bizonytalanságból adódó jelenségeken belül

Elemeztük a téves kivitelezések arányát a nyelvi tervezés szintjei szerint (3. ábra). A tervezés fogalmi szintjén egyik gyermeknél sem keletkeztek hibák, tehát az anyagban sem freudi elszólás, sem malapropizmus nem fordult elő. A gyermekeknél ugyanis – legalábbis egy bizonyos életkorig – nem beszélhetünk elfojtott elemről, mivel mindent kimondanak, ami eszükbe jut. A malapropizmus hiánya az óvodások mentális lexikonának sajátosságaival magyarázható. A fogalmi szint problémamentessége tehát a gyermek életkori sajátosságaiból adódik. A nyelvi tervezés szintjén a legmagasabb a hibaarány (41,0%). A leggyakrabban előforduló hibák között szerepelt az igekötő tévesztése (*lemelegedtem*), a névelő és az azt követő szó első hangjának téves összehangolása (*az kerianyuval*, *a autóval*), szám-személy egyeztetési hiba (*én pedig úgy csináltuk*). Ebben a korban már stabilizálódik a nyelvtani szabályrendszer, de a szabályhasználat még nem olyan automatikus, mint felnőttkorban. Az ebben az életkorban megjelenő hosszabb és összetettebb közlések esetén a szabályokat a gyermek nehezebben alkalmazza. A lexikális hozzáférés szintjén is viszonylag magas

a hibaarány (12,9% – *meg amit mondtam előbb, az a ő az a az a ő strand*), viszont egyáltalán nem találtunk hibát a fonológiai tervezés szintjén. Több kísérlet is azt mutatta, hogy ez a nyelvspecifikus tervezés relatíve stabil; például a magánhangzó-harmónia szabálya ellen a gyermek már 3 éves kora után nem vét, ez nagyon korán és erősen rögzülő nyelvspecifikus szabály (Gósy 2003). A vizsgált óvodások sok hibát ejtettek a lexikális hozzáférés és az artikulációs tervezés összehangolatlanságából adódóan (28,8%): a leggyakoribb jelenség az volt, amikor a gyermekek téves lexémát hívtak le a mentális lexikonból, majd félbehagyták azt, és másikat aktiváltak. Az artikulációs tervezés szintjén 15,1% a hibaarány, az anticipáció (*megnéztük a p- anyukám apukájával*) és a perszeveráció (*és felöltözöm... akkor egy tévét nézem, akkor felő, fe... bemegyünk a kocsiba*) nagyjából ugyanannyiszor jelentkezett. Metatézisre csupán három példát találtunk. Az artikulációs kivitelezés szintjén 2,2% a hibaarány: az egyszerű nyelvbottlás mindössze háromszor fordult elő a korpuszban (*ki se lehet riángatni a vízből*).



3. ábra

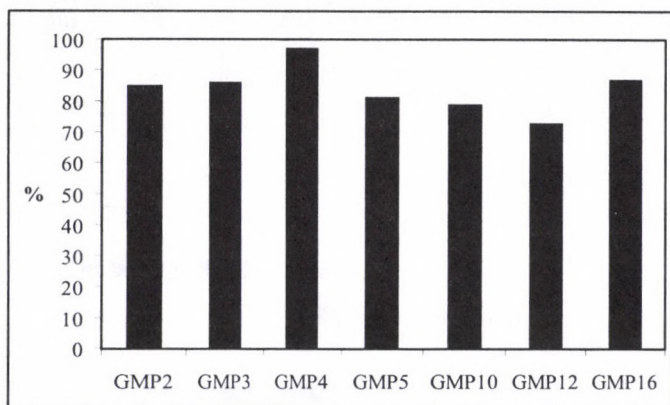
A téves kivitelezések aránya a tervezés szintjei szerint

Beszédészlelés és beszédmegértés

A jelen kutatásban hét részteszt eredményeit vizsgáltuk a GMP-diagnosztikából: öt az észlelési, kettő pedig a megértési teljesítményre vonatkozott. Elemeztük, hogy a 16 gyermek átlagosan milyen eredményeket produkált az egyes észlelési és megértési feladatokban (4. ábra).

Az óvodások a szűk frekvenciás mondatok észlelésében (GMP4) teljesítettek a legjobban (97%), bár nem mindegyik gyermek érte el az ebben az életkorban elvárható 100%-ot. Az észlelést vizsgáló tesztek közül a hangok sorrendiségének helyes azonosítása (GMP10) volt a legnehezebb a számukra (79%). Ez az átlag jelentősen elmarad az elvárható 100%-tól, vagyis az óvodások nagy része ezen a területen elmaradást mutat, amely később negatív hatással lehet az iskolai telje-

sítményre. A sorozatészlelés problémái ugyanis nagymértékben gátolják a mentális lexikon növekedésének lehetőségét, valamint megnehezítik az írás és olvasás elsajátítását. A szeriális észlelés zavart működése jellegzetes hibatípusokkal jár együtt: beszédhangok vagy szótagok kiesése, betoldása, esetleg felcserélése; a teljes hangsor torzítása (Gósy 2000, Imre 2006). A zajos mondatok (GMP2) és szavak (GMP3) azonosításában nagyon hasonló átlagértéket kaptunk (85%, ill. 86%). A zajjal fedett mondatok helyes ismétlésének átlaga megközelíti az elvárt értéket (90%), de a zajos szavak észlelésében a gyermekek jóval az elvárt 100% alatt teljesítettek.



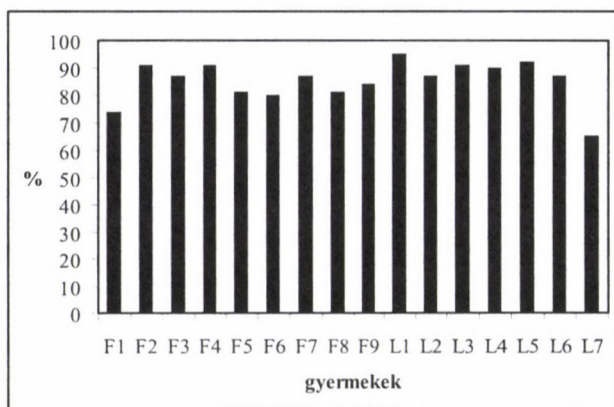
4. ábra

A GMP-résztesztek eredményeinek átlagértékei

A fonológiai észlelésre vonatkozó részteszt (GMP5) átlagos eredménye (81%) szintén elmarad az elérendő 90%-tól, vagyis ez a részfolyamat is több gyermeknél zavart működést mutat. Ez a részterület már komplexebb nyelvi feldolgozást végez, itt történik a fonémadöntés meghozatala.

A gyermekek az összes részteszt közül a szövegértésben (GMP12) teljesítettek a leggyengébben (73%), és itt jelentkeztek a legnagyobb egyéni különbségek is (20–90%). Az óvodások közül senki nem teljesítette a szövegértési feladatot hibátlanul, ahogy az ebben a korosztályban elvárható lenne. Az átlagos mondatértési teljesítmény ennél jóval magasabb (87%), bár ebben a résztesztben sem érte el minden gyermek az elvárt 100%-ot. Több száz iskolással elvégzett kutatás is igazolta, hogy a szövegértés eredményei minden korosztályban alatta maradnak a mondatértési teljesítménynek (Gósy 2000, Imre 2006). A szövegértés ugyanis nemcsak a szemantikai és szintaktikai szerkezetek értelmezését foglalja magába, hanem olyan bonyolultabb műveleteket is, mint a részletek azonosítása és rövid idejű tárolása, az ok-okozati viszony felismerése, a lényeg kiemelése, a tanulságok levonása stb.

Az egyes résztesztek eltérő folyamatműködéseket vizsgálnak, azonban az összehasonlíthatóság kedvéért a hét elemzett tesztben kapott eredményeket átlagolva minden óvodásnál meghatároztunk egy átlagos percepció teljesítményt (5. ábra).



5. ábra

A gyermekek percepció szintje

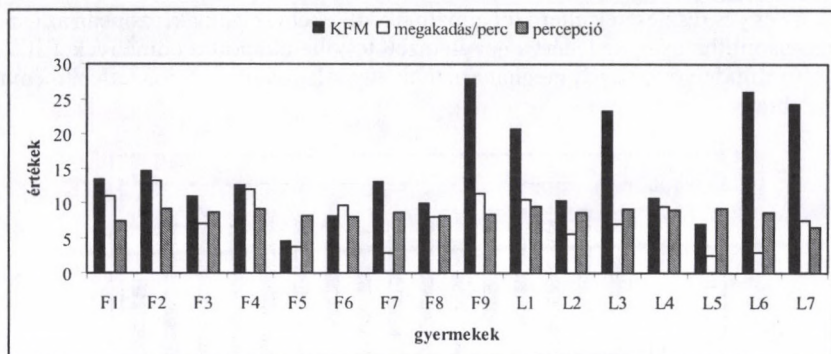
Az adatok szerint kettővel több lány érte el a 85% vagy a fölötti átlagot; tehát náluk kissé jobb teljesítményt mértünk, mint a fiúknál. A legjobb (95%) és a leggyengébb (65%) átlagot ezúttal lányok produkálták.

Két lány érte el a tizenhat óvodásból minden tesztben az ebben az életkorban elvárható értékeket (12,5%).

A beszédészlelés három – akusztikai, fonetikai és fonológiai – szintje között erős korreláció mutatható ki; vagyis e három részfolyamat egyikének zavara valószínűleg befolyásolja a többi működését is (akusztikai-fonetikai: $r = 0,658$; $p = 0,01$; akusztikai-fonológiai: $r = 0,657$; $p = 0,01$; fonetikai-fonológiai: $r = 0,543$; $p = 0,05$). Az akusztikai ($r = 0,543$; $p = 0,01$) és a fonetikai észlelés ($r = 0,799$; $p = 0,01$) elmaradott működése a szövegértési képességre is hatással van. Ha az akusztikai kulcsok felismerése és a fonetikai észlelés zavart mutat, akkor valószínűsíthető a gyermek hallásproblémája, és/vagy a grammatikai szerkezetek feldolgozási nehézsége. Ez utóbbiak pedig nagymértékben nehezítik az összefüggések felismerését, a lényeg kiemelését, a tanulságok levonását, vagyis a szöveg sikeres feldolgozását.

A produkció és a percepció összefüggései

Az egyének közötti legnagyobb eltérés a KFM-pontszámokban jelentkezett, míg a gyerekek között a percenként előforduló megakadások számában volt a legkisebb eltérés. A produkció és a percepció egyénenkénti összefüggéseit mutatja a 6. ábra.



6. ábra

A produkció és a percepció összefüggése egyénenként

(A függőleges tengelyen az „értékek” a KFM-pontszámot, a percnkénti megakadások számát és a GMP-tesztekben elért átlag tízzel osztott értékét mutatják egyénenként.)

Az ábrán látható értékek szemléltetik a Pearson-féle korrelációs együtttható adatait, miszerint hipotézisünkkel ellentétben nincs szignifikáns összefüggés a szerkezetek bonyolultsága, a megakadások és a percepció teljesítmény között. A jelen kutatás megerősítette a külföldi szakirodalomban (Yaruss–Newman–Flora 1999) megállapítottakat: a spontán beszéd szintaktikai komplexitásának mértéke nem előjelzője annak, hogy a közlést megszakítja-e valamilyen megakadásjelenség.

A magas KFM-pontszám járhat együtt kevesebb hibázással és gyengébb percepció átlaggal (L7), míg egy viszonylag alacsonyabb KFM-érték hasonló számú megakadással és feldolgozási teljesítménnyel párosulhat (F9). Arra is akadt példa, hogy az alacsonyabb KFM-érték (7,06) ellenére a percepció átlag (PÉ) a legjobbak között szerepel (L5). A közlések szerkesztettségének mértéke tehát nem utal szükségszerűen a percepció elmaradására, az életkornak megfelelően működő feldolgozási folyamatok pedig nem jelentik azt, hogy a gyermek szívesen beszél, és bonyolultabb szerkezetű közléseket valósít meg.

A produkció és a percepció összefüggését egy esetben sikerült statisztikailag igazolni: közepes korrelációt találtunk a gyermekek szeriális észlelése (GMP10) és KFM-pontszáma között ($r = 0,583$; $p = 0,05$). A sorrendiség észlelésének zavara tehát negatív hatással van a közlések szerkesztettségének mértékére, mivel alapvetően befolyásolja a mentális lexikon bővülésének lehetőségét, korlátozza az új lexémák elsajátítását és változatos használatát; és valószínűsíthetően összefüggésben van a munkamemória működésével is.

Az eredmények igazolták a tapasztalatot; függetlenül attól, hogy a gyermek szívesen beszél, hosszabb monologikus közléseket hoz létre, a feldolgozási folyamatai nem feltétlenül működnek az életkorának megfelelő szinten. A legmagasabb KFM-pontszámot elérő kisfiú (F1) például csak 40%-os teljesítményt ért

el a szövegértésben, vagyis nála ez a részfolyamat többéves elmaradást mutat. A szövegértés ilyen mértékű zavara később nagymértékben korlátozni fogja a tanulás eredményességét; noha a gyermek spontán beszédének komplexitása alapján nem volt feltételezhető percepciók elmaradása. Ugyanakkor az a kislány (L5), aki a legalacsonyabb KFM-pontszámot érte el, minden percepciók résztesztben életkorának megfelelően teljesített. Nem igazolódott tehát az a feltételezés, hogy ha a gyermek a monologikus közlések helyett csak egy-egy szóval vagy rövid szerkezettel vesz részt a „társalgásban”, akkor ez mindenképpen percepciók zavaráról árulkodik.

Következtetések

Az eredmények nagy egyéni különbségeket mutattak a gyermekek között a produkció és a percepció szintjében (a percenként előforduló megakadások számában nem találtunk ilyen nagy eltéréseket). A nagy individuális különbségek felvetik a kérdést, hogy az iskolában miként tudnak majd ezek a gyermekek együtt haladni az írott anyanyelv elsajátítása során.

Azt tapasztaltuk, hogy a gyermekek más stratégiákat használnak a tervezés diszharmóniáinak feloldására, mint a felnőttek. Az óvodásoknál az ismétlés a tervezés bizonytalanságának fő jelzője, és nem a hezitálás, mint a felnőtteknél. A kitöltött szünet stratégiáját a gyermekek – valószínűleg a felnőtt minta követésével – fokozatosan építik ki a közlésfolyamatok egyre bonyolultabbá válása során. A gyermekek számára ebben az életkorban még nagyobb gondot okoz a nyelvi tervezés előkészítése; a felnőtteknél a grammatikai hibák csupán 6,5%-át adják az összes megakadásnak (Gósy 2003), az óvodásoknál ennek kétszeresét.

A percepciók folyamatokat vizsgálva az eredmények azt mutatták, hogy az észlelés akusztikai, fonetikai és fonológiai szintje szoros összefüggést mutat, így az egyik folyamat zavara negatív hatással van a többire is. A nem ép működés hatással van az artikuláció pontosságára, a szókincs bővülésére, a morfológiai struktúrák felismerésére; amely később gátolja az írott anyanyelv és az idegen nyelvek elsajátítását, vagy például a memoriterek sikeres tanulását (Gósy 2000).

A KFM-érték és a szeriális észlelés összefüggését a statisztikai elemzés igazolta: a sorrendiség zavara negatív hatással van a közlések szerkesztettségére, mivel befolyásolja többek között a szókincs bővülésének lehetőségét. Az észlelési tesztek közül a gyermekek a szerialitást mérő feladatban teljesítettek a leggyengébben. A szerialitás problémája abban is megjelenik, hogy az óvodások négyszer annyi sorrendiségi hibát ejtettek a spontán beszédben, mint a felnőtt adatközlők (Gósy 2003).

Hipotézisünkkel ellentétben az adatok azt mutatták, hogy az óvodásoknál nincs szoros összefüggés a produkció és a percepció között. Egy jó szintű beszédprodukció mögött tehát akár komoly percepciók zavar is meghúzódhat. Ezért mindenképpen fontos lenne a feldolgozási folyamatok tesztelése a gyermekeknél bizonyos rizikótényezők megléte esetén, illetve kivétel nélkül iskolába lépés előtt. Természetesen az is hátrányt jelenthet a gyermek számára az isko-

lában, ha a megfelelő percepció szint mellett nehezen fejezi ki magát, közlései kevésbé szerkesztettek és nyelvtanilag hibásak. Ennek megfelelően a közlésegség fejldési mutatójának alacsony értéke is eljelezhet iskolai nehézségeket.

Irodalom

- Beke Anna dr. 1996. A beszédpercepció fejlődésének neurológiai háttere. In Gósy Mária (szerk.): *Gyermeckori beszédészlelési és beszédmegértési zavarok*. Nikol Kkt., Budapest, 32–53.
- Gerebenné Várbíró Katalin – Gósy Mária – Laczkó Mária 1992. *Spontán beszéd-megnyilvánulások szintaktikai elemzése DDS technika segítségével*. Kézirat. Budapest.
- Gósy Mária 2000. *A hallástól a tanulásig*. Nikol Kkt., Budapest.
- Gósy Mária 2003. A spontán beszédben előforduló megakadási jelenségek gyakorisága és összefüggései. *Magyar Nyelvőr* 127. 257–277.
- Gósy Mária 2004. A spontán magyar beszéd megakadásainak hallás alapú gyűjteménye. *Beszéd kutatás 2004*. 6–18.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2006. *GMP-diagnosztika*. Nikol Kkt., Budapest.
- Horváth Viktória 2004. Megakadási jelenségek a párbeszédekben. *Beszéd kutatás 2004*. 187–199.
- Imre Angéla 2006. Az olvasászavar és a beszédfeldolgozási folyamatok összefüggései. *Beszéd kutatás 2006*. 160–171.
- McLaughlin, Scott F. – Cullinan, Walter, L. 1989. Disfluencies, utterance length, and linguistic complexity in nonstuttering children. *Journal of Fluency Disorders* 14/1. 17–36.
- Papp István 2006. *Male brain – female brain*. Előadás. Pszicholingvisztikai Nyári Egyetem, Balatonalmádi.
- Schneider Júlia – Simon Ferenc 1996. Iskolai kudarcok beszédpercepciósi hátteréről. In Gósy Mária (szerk.): *Gyermeckori beszédészlelési és beszédmegértési zavarok*. Nikol Kkt., Budapest.
- Yaruss, J. Scott – Newman, Robyn M. – Flora, Tracy 1999. Language and disfluency in nonstuttering children’s conversational speech. *Journal of Fluency Disorders* 24/3. 185–207.

A tanulmány az OTKA T046455 számú pályázat keretében készült.

A BESZÉDHANGOK MEGKÜLÖNBÖZTETÉSÉNEK FEJLŐDÉSE

Gósy Mária

Bevezetés

A beszédészlelés a nyelvi információt tartalmazó akusztikai jelsorozat feldolgozásával indul. Ennek kezdete az elsődleges hallási elemzés, ahol néhány előzetes döntés történik a bejövő jelsorozat akusztikai sajátosságait illetően. Ezek a döntések nagyjából behatárolják a felfogott (észlelt) hullámforma tulajdonságait; felismerjük például, hogy zene vagy beszéd hangzott-e el, énekszót vagy hangszert hallottunk, gyors volt-e vagy lassú, magas vagy mély, halk vagy erős (egyszersmind egyéb szubjektív ítéleteket is alkothatunk). A hierarchikusan egymásra épülő akusztikai, fonetikai és fonológiai szintek működését további beszédészlelési részfolyamatok egészítik ki (Gósy 2005). A különböző beszédészlelési folyamatok csaknem egyidejűleg működnek, és egymással igen bonyolult módon tartanak kapcsolatot. A beszédhang-differenciálás egyfelől a fonetikai, másfelől a fonológiai szint működését egészíti ki.

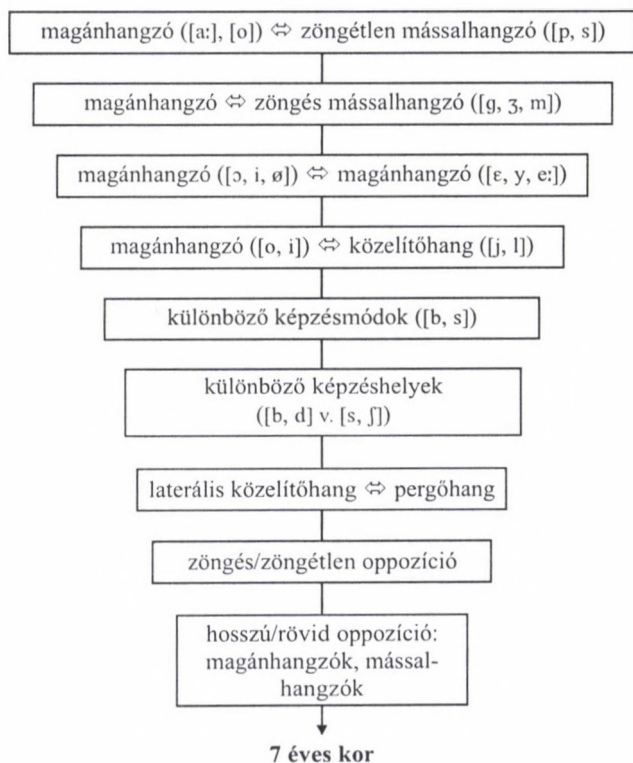
Az anyanyelv-elsajátítás folyamatában univerzális jelenség, hogy a gyermek mindig a nagyobb különbségek észlelésétől halad a kisebbek felé (Grunwell 1987). A beszédhangok tekintetében ez azt jelenti, hogy legelőször a magánhangzókat és a mássalhangzókat képes elkülöníteni egymástól, illetőleg azokon belül is a nagy artikulációs különbségekből fakadó nagy akusztikai következménnyel járó eltéréseket észleli. Ilyen például egy magánhangzó és egy felpattanó zöngétlen zárhang különbsége. Majd fokozatosan válik a gyermek képessé az egyre kisebb eltéréssel realizálódó beszédhangok differenciálására. A folyamat ezen belül nyelvspecifikus, hiszen aszerint történik a nyelvi diszkriminációs fejlődés, hogy milyen fonetikai sajátosságai vannak az egyes anyanyelvi fonémarealizációknak. Az individuális különbségek mintegy hatéves korig meglehetősen nagyok lehetnek ép fejlődés esetén is. A gögicselés hangjai fonetikailag különböznek, ami arra (is) utal, hogy a csecsemő észleli a környezetében hallható hangok sokféleségét. Ezeknek azonban nincs nyelvi funkciójuk. Az első szóhangsorokban, illetőleg holofrázisokban megjelenő beszédhangok már bizonyos fonémakezdeményeknek tekinthetők, hiszen egyfajta nyelvi realitással bírnak, és előkészítik a valódi fonémarealizációk megjelenését. Ehhez azonban az szükséges, hogy a kisgyermek képes legyen a feldolgozott hangsorokban a beszédhangokat fonémáknak megfeleltetni. Voltaképpen ennek tükröződése a holofrázisok fonetikai szerkezetének alakulása. A gyermek első tartalmas szavainak alakja többé-kevésbé felidézi a felnőtt nyelvi mintát; az artikulációs és nyilvánvaló percepci-

ós korlátok azonban nem teszik lehetővé a gyermek számára a tökéletes ismétlést. A kezdetekben ennek következtében a hallott hangsorok egy részlete jelenik meg a gyermeki holofrázisokban. Egy átlagos fejlődésű, 15 hónapos gyermek első 18 szavának elemzése azt mutatja, hogy képes az alsó nyelvállású veláris és palatális magánhangzók megkülönböztetésére (*nem* és *ham*), sőt a bilabiális és alveoláris zöngés felpattanó zárhangokéra is (*ada* 'adjál' és *aba* 'labda'). Nem állítható, hogy ezek a beszédhangok már fonémaértékűek lennének, de feltétlenül valamiféle stabilizálódásról és nyelvi rendszerről tanúskodnak. Az utóbbi tízenöt év kutatásai megerősíteni látszanak azt a feltételezést, hogy a gyermek szókinsének kiépülése nemcsak az artikulációs és akusztikus lenyomatok, hanem a fonológiai leképeződések révén alakul ki (Ingram 2001). Ennek a hipotézisnek a hátterében az a felfogás áll, hogy a gyermek beszédészlelése révén kialakult „szókins” a percepcióis képességgel együtt elegendő ahhoz, hogy megkezdődjön a fonológiai szerveződés az első szavak létrehozásához. A holofrázisok nem mindig stabilak a fonetikai szerkezetüket tekintve. A 'nem' szó alakvariánsai például egy év körüli gyermek ejtésében a következők: *mem*, *nem*, *nen*, *mam*, *m[æ]m*, avagy a 'kutya' jelentésű *vau* hangsor produkciói: *vúvú*, *vauí*, *[w]auí*, *[w]au*, *[w]u[w]u*. A szóhangsorokban található magánhangzók egyikének artikulációja sem biztos, az első esetben még a nyelv vízszintes mozgása szerint sem egyértelműen differenciálódott ez a magánhangzó. A nazálisok pedig gyakorlatilag egymás variánsai, egymással cserélődnek ugyanabban a hangsorban a jelentés változása nélkül. A másik példában ugyancsak bizonytalan a magánhangzó, amely monoftongusként és diftongusszerű elemként egyaránt realizálódik, időtartama sem stabil; továbbá a mássalhangzó képzési helye is változó. Ezekben a példákban tehát nyilvánvalóan olyan beszédhangejtés történik, amelynek hátterében nem határozható meg fonéma. A sokasodó szóhangsorok egyfelől stabilizálják az artikulációs mozgásokat, másfelől biztosítják a kialakuló fonémák realizációinak azonosságát. A magyarban a beszédhang-megkülönböztetés fejlődése az alábbiak mentén történik; gyakori a párhuzamosság, illetőleg az átfedés a folyamatban (1. ábra).

A gyermekek diszkriminációs képességét eseményhez kötött agyi potenciálok (EKP-k) mérésével is elemezték. Azt találták, hogy az EN (eltérési negativitás) szerint – amely az EKP-k egy kicsiny komponense – a fonémamegkülönböztetés tekintetében a változások ugrásszerűek az első és a második osztályosok között, majd a különbségek csökkenő tendenciát mutatnak (Csépe 2003). Ez az eredmény azzal is magyarázható, hogy az olvasás, írás nagymértékben hozzájárul a gyerekek helyes beszédhang-diszkriminációs működéséhez.

Jól ismert, hogy a gyermekek között, főként fiatalabb életkorban, nagy különbségek tapasztalhatók az anyanyelvi ismeretekben, az egyes elsajátítási szakaszok megjelenésében és tartamában (Smith–Kenney 1998). Hatéves korra elvárható, hogy egy egynyelvű gyermek az anyanyelvi beszédhangjait képes legyen megkülönböztetni; ez alól a hosszú-rövid oppozíció pontos észlelése még kivétel (lásd később). A pontos megkülönböztetés hiányában ugyanis egyéb be-

szédészlelési nehézségek, valamint a szókincs életkor-specifikus növekedésének zavarai is várhatók. A diszkriminációs képesség nem megfelelő életkori szintje hatással van az írott anyanyelv, az olvasás, írás megtanulására. Az anyanyelv-fejlődés során tapasztalt megkülönböztetési hibákat nem valamiféle percepciósi zavarra, hanem a kialakulatlan artikulációra vezetik vissza a kísérleti eredmények alapján angol anyanyelvű gyermekeknél (Winitz–Bellerose 1968). A beszédhibásoknál összefüggést találtak a diszkriminációs mint észlelési zavar és a beszédhiba között. Saját korábbi vizsgálataink igazolták, hogy a tanulási nehézségeik miatt vizsgált gyermekek egy részénél nem működik tökéletesen a beszédpercepciósi mechanizmusnak ez a részfolyamata (Gósy 2000). Noha a beszédhangok megkülönböztetésének fejlődési folyamata elméletileg nyilvánvaló egy nyelv hangrendszerének ismeretében, éppen az egyéni különbségek miatt mégis szükségesek a pontos, objektív adatokra támaszkodó ismeretek.



1. ábra

A beszédhangok differenciálásának fejlődése magyar gyermekeknél

Kísérletsorozatot terveztünk azzal a céllal, hogy megtudjuk, vajon az univerzálisnak tekinthető tendenciát mennyiben támasztják alá a kísérleti adatok. Hipotézisünk az volt, hogy fiatalabb óvodás korban az elméletileg fejlettebb észlelési működést követelő megkülönböztetések még nehézségekbe ütköznek, első osztályos kortól azonban már csak véletlenszerűen fordulnak elő hibázások. Azt feltételeztük, hogy például a zöngéesség megítélése a kisebb gyermekek esetében nem lesz problémamentes; a nagyobbak azonban legfeljebb a beszédhangok időtartamának megítélésében tévednek. A hipotézisünk háttérében az az artikulációs tény áll, hogy míg a felnőtt nyelvi környezetben a beszédhangok időtartamának realizálása meglehetősen bizonytalan, addig például a zöngéesség érvényesítése jóval egyértelműbb.

Kísérleti személyek, tesztanyag, módszer

Összesen 450 gyermeket teszteltünk négy korcsoportban: 5 éveseket (közép-sős óvodásokat), 6 éveseket (nagy csoportos óvodásokat), 7, 8 és 9 éveseket (első, második és harmadik osztályosokat). Az öt évesek száma 100, a hat éveseké 200, a hét-, nyolc- és kilenc éveseké pedig 50-50 volt. Minden életkori csoport fele lány, fele fiú. A gyermekeket véletlenszerűen válogattuk ki fővárosi és vidéki városok (kisebb és nagyobb települések) óvodáiból és iskoláiból. Ép halló és ép beszédű gyermekek voltak (az artikuláció tisztaságát fokozottan vettük tekintetbe, hogy ne legyen hatással a megkülönböztetés esetleges zavarára).

A GMP-diagnosztika (Gósy 2006) 17-es altesztjét használtuk, amely 23 egy, két és három szótagú értelmetlen hangsort tartalmaz. 15 hangsor egymástól egyetlen képzési jegyben különbözik (pl.: *azsá ~ asá, mőz ~ nőz, nőcs ~ nős, nazirő ~ nazilő*); 8 pedig teljesen egyforma (pl.: *fész ~ fész*). A magánhangzók az időtartam, a mássalhangzók a képzési mód, a képzési hely, a zöngéesség és az időtartam tekintetében különböznek egymástól. A logatomok random sorrendben következnek egymás után. A kísérletvezető ismertette a gyermekkel a feladatot, majd a sikeres próbát követően futtatta a tesztet, minden gyermekkel egyénileg, csendesített szobában. A gyermeknek egyszeri elhangzás alapján kellett döntenie arról, hogy vajon egyforma, avagy nem egyforma volt-e a hallott két hangsor. A kísérletvezető feljegyezte, ha a gyermek hosszan gondolkodott, vagy ha hangosan kiejtette a teszthangsorokat (ez utóbbi a nagyfokú döntési bizonytalanság jele). Az eredményeket a helyes döntés szerint összesítettük, külön-külön az eredetileg azonos és az eredetileg különbözőeket, majd az adatokat statisztikailag elemeztük (ANOVA, SPSS 8.0 program használata, 95%-os szinten).

Eredmények

Az adatok összesítése jellegzetes különbségeket mutatott az egyes korcsoportok között egyetlen kivétellel. A hat- és hét évesek teljesítménye között nem volt szignifikáns különbség. Az életkor növekedésével javuló teljesítményt csak kis mértékben tapasztaltuk, és még a harmadik osztályosoknál sem találtunk sok hibátlan megoldást. Elgondolkodtató ugyanakkor, hogy az öt évesek 8%-ának

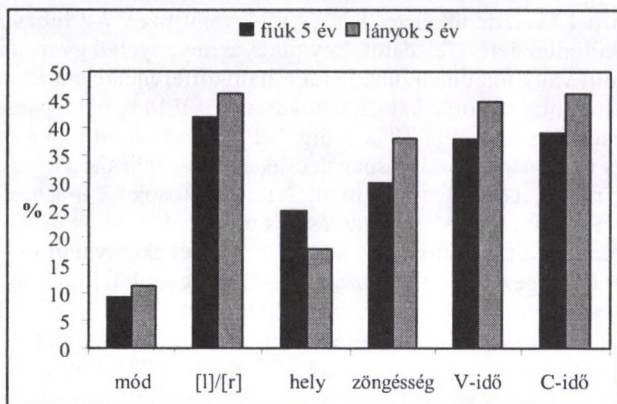
(nyolc gyermek) tesztje tökéletes lett, és több mint tízen 1-2 hibával oldották meg a megkülönböztetési feladatot. Egy magyar anyanyelvű gyermek tehát képes hibátlanul vagy majdnem hibátlanul a hangdifferenciálásra az adott teszt-helyzetben már ötéves korában is. A hatévesek közül 14%-nál tapasztaltunk hibátlan teljesítményt, további 18% pedig 1-2 hibával oldotta meg a feladatot. Ugyanakkor a harmadik osztályosoknak csak a 40%-a teljesített tökéletesen, noha már több mint két éve olvasnak, írnak. A másodikosok 42%-a, az elsősöknek pedig 22%-a oldotta meg a tesztet tévesztés nélkül.

Az adatok átlagait, a minimum- és maximumértékeket, valamint az átlagtól való eltérést összegeztük az 1. táblázatban. A legnagyobb fejlődés az öt- és hatévesek, valamint a hét- és nyolcévesek között tapasztalható. A statisztikai adatok azt mutatják, hogy szignifikáns változás van az egyes korcsoportok között (a Wilcoxon-próba szerinti rangszám: $-33,764$, $p = 0,0002$). Elemeztük a kapott adatokat aszerint, hogy vajon a gyermekek képesek-e észlelni a beszédhangok közötti egyetlen képzési jegyben jelentkező különbségeket a zöngéesség, a képzésmód és képzéshely, valamint az időtartam szempontjából.

1. táblázat: A helyes beszédhang-diszkrimináció átlageredményei

| Csoportok | Helyes megkülönböztetés (%) | | | |
|-----------|-----------------------------|-----------------|---------|---------|
| | Átlag | Átlagos eltérés | Minimum | Maximum |
| 5 évesek | 77,37 | 19,56 | 0,0 | 100 |
| 6 évesek | 82,95 | 21,80 | 21,7 | 100 |
| 7 évesek | 82,70 | 17,15 | 34,7 | 100 |
| 8 évesek | 93,13 | 8,87 | 65,2 | 100 |
| 9 évesek | 91,30 | 11,45 | 52,1 | 100 |

Az ötévesek legnagyobb mértékben a mássalhangzók időtartamának elkülönítésében tévedtek; itt a helyes megoldás mindössze 15%. Valamivel jobb eredményt kaptunk a magánhangzók időtartamának pontos megkülönböztetésére, ebben a helyes észlelés átlaga 28% volt. Ötéves korban még komoly gondot okoz a zöngéesség tényének megítélése, a helyes differenciálás átlaga 49% lett. A képzési hely szerinti elkülönítés már jóval könnyebb feladatnak mutatkozott, a helyes megoldások aránya 78,5% lett. Legkönnyebb a képzési mód szerinti elkülönítés volt, a helyes teljesítmény átlaga 90,6%. A fonetikai osztályozás szerint a laterális approximáns és a pergőhang képzési módban különböznek, mégis célravezetőbb volt ezt a két mássalhangzót külön kategóriaként elemezni. Ennek az volt az oka, hogy mindkettőt a képzési módnál tárgyalva, hamis képet kaptunk volna a differenciálási teljesítményről. Valószínűsíthetően a gyermekek saját artikulációs korlátainál fogva, továbbá a két beszédhang akusztikai sajátosságai miatt is, a [r] és a [l] mássalhangzók megkülönböztetése hosszabb időn keresztül meglehetősen nehéz a számukra. Noha statisztikailag igazolható különbség nincs a lányok és a fiúk teljesítménye között, a képzési hely kivételével a fiúk kevesebbet hibáztak (2. ábra).



2. ábra

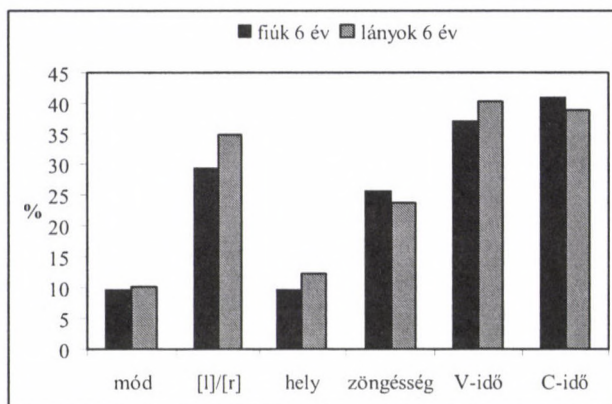
A beszédhang-differenciálás tévesztései ötéveseknél

Elemeztük, hogy az azonos hangsorok azonosságának megítélésében tévedtek-e a gyermekek. Azt találtuk, hogy a tévesztések aránya kicsi, mindössze 8%-ban hibáztak a fiúk és 8,2%-ban a lányok, vagyis ekkor különbözőnek ítélték az azonos hangsorokat. Ez a differenciálási készség bizonytalanságát mutatja.

A hatéveseknél elkülönítettük a fiatalabb (6;6 alatti) és az idősebb (6;6 fölötti) gyermekek teljesítményét is. Azt találtunk, hogy a fiatalabb fiúk jobban teljesítettek az idősebeknél, bár a különbség kicsi, nem egészen 10%; a lányoknál az 5%-nyi különbség ugyanakkor az idősebbek javára jelentkezett. Sem a kor, sem a nem tekintetében nincs azonban statisztikailag igazolható különbség a gyermekek között. Mindkét korcsoportban voltak olyan gyermekek, akik hibátlanul oldották meg a feladatot (19-en a fiúk és 8-an a lányok között). A maximális hibaszám a fiatalabb fiúknál 11, a nagyobbaknál 15, a fiatalabb lányoknál 17, az idősebeknél pedig 16 volt. A két korcsoportot tekintve gyakorlatilag nincs különbség a tekintetben, hogy mely képzési jegyek eltérését volt nehezebb felismerniük a gyermekeknek (3. ábra). A hatéves lányok a zöngéesség és a mássalhangzók időtartamának eldöntésében valamivel jobban teljesítettek, mint a hatéves fiúk.

Az öt- és hatévesek beszédhang-megkülönböztetési eredményeit összehasonlítva, jól látható a fejlődés. Legfeltűnőbb a képzési hely tekintetében, valamint a laterális approximáns és a perghang megkülönböztetésének pontosságában. Nagyon csekély ugyanakkor az eltérés a magán- és a mássalhangzók időzítési viszonyait tekintve. Meghökkenítő volt a hatéves óvodások és a hétéves első osztályosok közötti – már említett – teljesítményazonosság. Azt feltételeztük, hogy hétéves korban nemcsak – akár ugrásszerű – fejlődést fogunk tapasztalni, de mindenképpen különbséget látunk az óvodásokhoz képest, hiszen ekkor már az olvasás és írás hatásával is számolhatunk. Az adatok azonban csak a minimumértékben mutatnak tíz százalék körüli „javulást”. Hasonlóképpen nem várt ered-

mény a második és harmadik osztályosok beszédhang-differenciálási teljesítménye. Noha statisztikai különbség nincs közöttük, az átlagérték és a minimumérték a másodikosoknál jobb. Elgondolkodtató, hogy még a harmadik osztályosok között is akadnak igen gyenge teljesítmények, amelyek felvetik az olvasás- és írástanítás módszertani kérdéseit is.

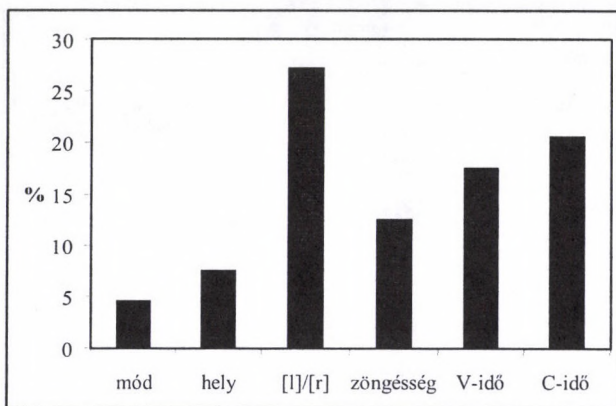


3. ábra

A beszédhang-differenciálás tévesztései hatéveseknél

Minden korcsoportban akadtak olyan gyermekek, akik csak igen hosszú reakcióidővel voltak képesek – a jó vagy hibás – döntésüket meghozni. A hosszabb gondolkodás azonban nem feltétlenül eredményezett helyes beszédhang-megkülönböztetést, sőt az esetek nagyobb részében ekkor hibáztak a gyermekek. Már öt-hat éves korra elvárható, hogy ebben a tesztben a gyermekek egyszeri hallás után azonnal, különösebb gondolkodás nélkül, mintegy „automatikusan” képesek legyenek a megkülönböztetésre. A gondolkodás bizonytalanságot, avagy diszkriminációs nehézséget jelent, és még jó döntés esetén is negatívan befolyásolja például a gyermekek írását és helyesírását. A kísérletben részt vevők mintegy harmadánál tapasztaltuk, hogy megpróbálták hangosan megismételni az elhangzott hangsorokat, és így segíteni önmagukat a döntésben. Noha minden ilyen esetben megkértük a gyermeket, hogy ne ejtse ki a hallottakat, gyakran nem voltak képesek a felszólításnak eleget tenni. Különösen az iskolásokra volt jellemző az ismételt auditív csatorna felhasználási kísérlete. Ez egyfelől jelzi ugyan a bizonytalanságot, másfelől viszont egy kiegészítő stratégia működtetését mutatja; a gyermek ösztönösen érzi, hogy az ismételt hangoztatás segíthet neki a differenciálásban. A 4. ábra szemlélteti az egyes képzési jegyeknek megfelelő téves megkülönböztetés átlagait az összes tesztelt iskolás gyermeknél. Az óvodásokhoz képest határozott javulás tapasztalható az iskolásoknál, kivéve a [r]/[l] megkülönböztetését, ebben gyakorlatilag nincs változás. Az elvárt, vagy-

is hibátlan hangdifferenciálást azonban a tesztelt iskolások egyetlen képzési jegy esetében sem érték el. Ezt annak tükrében kell tekintenünk, hogy ezek a gyermekek már írnak és olvasnak. A statisztikai elemzések szignifikáns különbséget mutattak az egyes képzési jegyek alapján történt megkülönböztetések között (Kruskal–Wallis-próba: 1836,212; $p = 0,0004$).

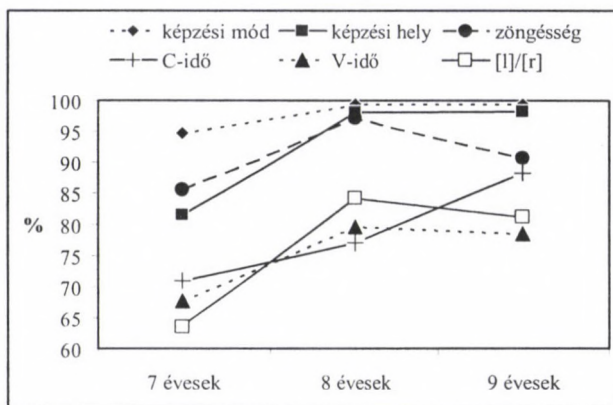


4. ábra

A beszédhang-differenciálás tévesztései iskolásoknál

Valamennyi grafikon követi a diszkriminációs fejlődésben az anyanyelv-elsajátítás elméleti tendenciáját; ha azonban a módhoz hozzávesszük a [l]/[r] megkülönböztetést is, akkor a képzési mód mint jegy esetében gyengébb teljesítményt kapunk. A [l]/[r] mássalhangzók megkülönböztetési nehézsége – mint láttuk – gyakran még hétéves kor fölött is jellemzően fennáll. Az 5. ábrán az eredményeket az iskolások korcsoportjai szerint részleteztük.

Várhatóan az időtartam-megkülönböztetések okozzák a legnagyobb problémát, annak ellenére, hogy a magyarban mind a magánhangzók, mind a mássalhangzók tekintetében az időtartam fonológiailag is releváns tényező. Ez azt jelenti, hogy azonos hangsorokban az időtartambeli eltérések jelentéskülönbségeket eredményeznek (pl. *irat* ~ *írat*, *örül* ~ *őrül*, *halom* ~ *hallom*, *kasza* ~ *kassza*). Adataink szerint a mássalhangzók időzítési különbségei 9 éves korban kissé jobbakká válnak a magánhangzókénál (fiatalabb gyermekeknél ezt nem tapasztaltuk), de egyik tekintetében sem teljesítenek a gyermekek tökéletesen. A mássalhangzók időtartamában a helyes megkülönböztetés fejlődése nagy valószínűséggel az írott anyanyelv tanításával van összefüggésben. A képzési mód, hely és a zöngéesség tényezőjének felhasználása nyolcéves korra megközelíti a kívánt szintet, a kilencéveseknél azonban tendenciaszerűen gyengébbek az eredmények. Az iskoláskor kezdetén a gyermekek jó része nagy elmaradást mutat az elvárt teljesítménytől.



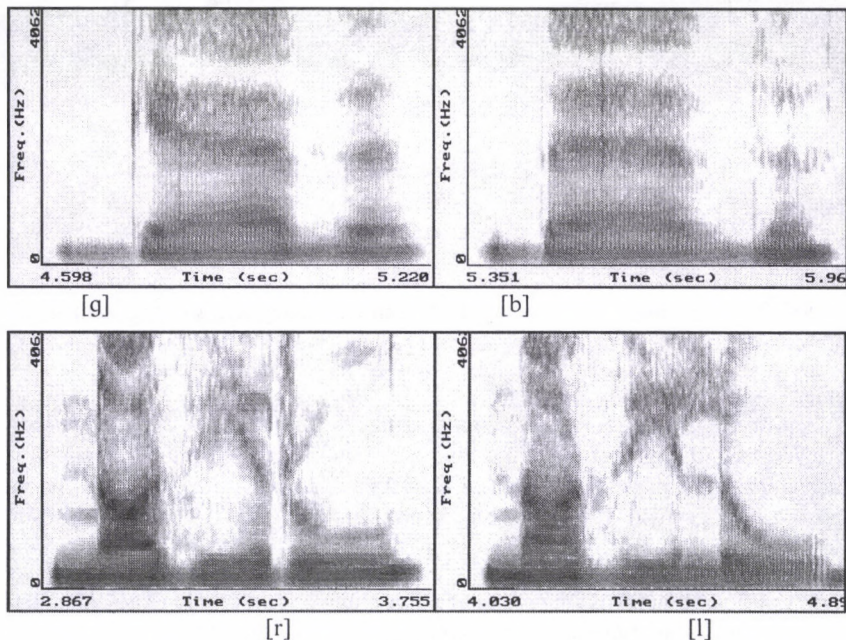
5. ábra

A helyes differenciálási teljesítmény a képzési sajátosságok függvényében iskolásoknál

Feltételeztük, hogy az akusztikusan kisebb különbséget mutató beszédhangok megkülönböztetése fog nehézséget okozni a gyermekeknek, az akusztikailag nagy(obb) eltéréssel jellemezhetők viszont nem. Az adatok nem igazolták ezt az összefüggést. A 6. ábra hangszíneképei a *gev ~ bev* és a *nazirú ~ nazilú* hangsor-párokról készült spektrogramokat szemléltetik (4000 Hz-es frekvenciatartományban, CSL 4300B típusú digitális jelfeldolgozóval). A felső ablakpárban a hangsorkezdő mássalhangzók akusztikai különbsége meglehetősen kicsi, hiszen a két zöngés zárhang között csupán a képzés helyében van eltérés. A bal oldalon látható *gev* hangsor első mássalhangzójára változóan intenzív, a 4000 Hz-es tartományban megjelenő zár-felpattanás jellemző, az ezt követő magánhangzó F_2 -je mintegy 200 Hz-es csökkenést mutat a tiszta fázisig. A jobb oldali ablakban látható [b] zár-felpattanása kevésbé intenzív, és nem jelenik meg a teljes (látható) frekvenciatartományban; az öt követő magánhangzó második formánsa kismértékű emelkedést mutat. A két mássalhangzó közötti akusztikai különbség tehát kicsinek mondható.

Az alsó ablakpár bal oldalán a *nazirú* hangsor pergőhangja jellegzetes akusztikai szerkezetű, két pergetés lenyomata látszik a hangszíneképen, a mássalhangzó időtartama mintegy 70 ms. A jobb oldali ablakban a *nazilú* hangsor spektrogramja látható, az ugyanazon magánhangzó-környezetben kiejtett közelítőhang időtartama több mint kétszerese a pergőhangénak. Akusztikai szerkezete magánhangzósrzerű, formánsai jól meghatározhatók. Bár az adott pergőhang és a közelítőhang csupán képzési módjukban térnek el, hiszen artikulációs helyük és zöngés voltuk azonos; az akusztikai szerkezetük ugyanakkor nagy különbséget mutat. Mind a frekvencia, az intenzitás és az időtartam paraméterei is jellegzetesen eltérnek. A gyermekek diszkriminációs teljesítménye mégsem az akusztikai kü-

lönbség szerint alakul. A *gev* ~ *bev* hangsorpárt 93,25%-ban (a 6 évesek 87%-ban, a 7 évesek 86, a 8-9 évesek 100-100%-ban), míg a *nazirü* ~ *nazilü* hangsorokat csak 68,25%-ban (a 6 évesek 44%-ban, a 7 évesek 63,5, a 8 évesek 84,3, a 9 évesek 81,2%-ban) különböztették meg helyesen. Ez felveti azt a lehetőséget, hogy a gyermekek diszkriminációja nemcsak a beszédhangok akusztikai különbségének, hanem saját artikulációs mozgásaiknak is függvénye, ami alátámasztja a beszédészlelésnek az artikulációs mozgásokkal feltételezett szoros kapcsolatát.



6. ábra

A *gev* ~ *bev* (fent) és a *nazirü* ~ *nazilü* (lent) hangsorok hangszínképei

Nem feltételeztük, hogy az azonos hangsorok felismerése nehézséget jelent az iskolásoknak. Az eredmények szerint azonban az azonos logatomokat az iskolás gyermekek 5,4%-ban tévesztették, azaz különbözőnek ítélték (2. táblázat).

Az azonos hangsorok különbözőnek ítéltése főként azoknál a gyermekeknél fordult elő, akiknek a megkülönböztetési teljesítménye egyébként is gyenge volt. Az óvodások bizonytalansága e tekintetben kissé nagyobb mértékű volt (8% körüli).

A számos hibátlan, illetve elfogadhatóan jó teljesítmény mellett egy 8 éves fiú a 23 teszthangsor-pár azonosságát, illetve különbözőségét mindössze 52%-ban döntötte el helyesen. Három esetben egyáltalán nem adott választ. Valamennyi

képzésijegy-típus tekintetében tévedett, az időtartam-különbségek felismerésére gyakorlatilag képtelen. Hasonlóan gyenge eredményeket tapasztaltunk egy 9 éves kislány esetében. Diszkriminációs észlelése igen gyenge, a hallott hangsorokat igyekezett gyakran hangosan megismételni. Átlagos teljesítménye 58%, hibázásai valamennyi képzésijegy-típust érintik, de nem azonos mértékben. Jellegzetesen téved az azonos hangsorokról adott döntéseiben is. Reakcióideje időnként erősen megnövekszik. A két iskolás gyermek teljesítménye jelentősen gyengébb, mint az öt évesek átlaga.

2. táblázat: Az azonos hangsorok helyes megkülönböztetésének átlagai iskolásoknál

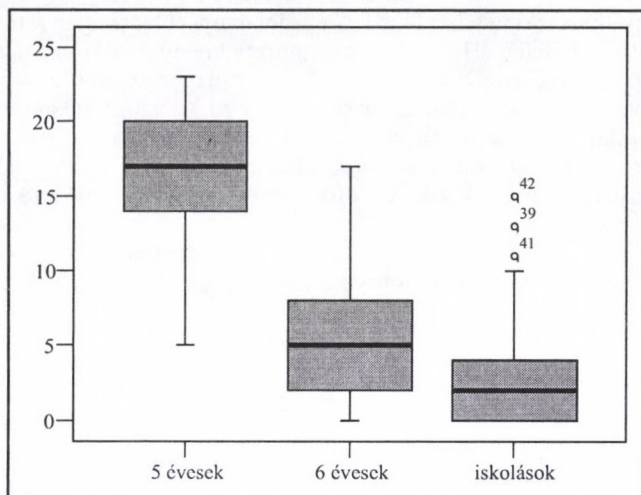
| Adatok (%) | Életkor | | |
|-----------------|----------|----------|----------|
| | 7 évesek | 8 évesek | 9 évesek |
| Átlag | 94,90 | 98,10 | 94,50 |
| Átlagos eltérés | 1,52 | 3,92 | 5,92 |
| Minimumérték | 93,48 | 88,00 | 82,76 |
| Maximumérték | 97,83 | 100,00 | 100,00 |

Következtetések

A jelen kutatás első ízben vizsgálta nagyobb gyermekcsoportok részvételével a beszédhang-megkülönböztetés fejlődését öt éves kortól kilencéves korig. Az elméletileg feltételezett fejlődést nem teljes egészében igazolták az adatok (vö. 7. ábra). Nem állíthatjuk, hogy első osztályos kortól már csak véletlenszerűen fordulnának elő hibázások. A differenciálás még kilencéves korban sem működik százszázalékosan, sőt mintha némi fejlődési megtorpanást mutatnának a kapott adatok.

Ez kétféleképpen interpretálható: vagy úgy, hogy a gyermekek ezen perceptív folyamatok nem működik tökéletesen, vagy úgy, hogy a magyar gyermekek differenciálási képessége tízéves korra válik tökéletessé. Ez utóbbi esetben azonban az írott anyanyelv tanításában ennek a ténynek feltétlenül meg kell jelennie. Megítélésem szerint inkább az első feltételezés a jó; ezt az 5-6 évesek között tapasztalható jó differenciálási teljesítmények igazolják egyfelől, illetőleg az, hogy rendszerint a zavarokat mutató gyermekek egyben olvasási, helyesírási nehézséggel küzdenek.

A perceptív nehézségeket mutató gyermekek teljesítményét a nagyothallókéhoz hasonlítják (Bishop 1992). A lassabb feldolgozás, a perceptív zavar esetén a gyermek nem tudja összevetni a bejövő akusztikai információt a tárolt ismeretekkel, és ez gátolja a nyelvi sajátosságok felismerését (Evans 2001). Jakobson azt állította, hogy könnyebb a beszédhangokat megkülönböztetni, mint azonosítani (vö. Hörmann 1971: 67); ezt az állítást kísérletek is megerősítették (Gósy 2005).



7. ábra

Az összes beszédhang-differenciálási tévesztés átlaga és szórása
(a hibák számában megadva)

A jelen kísérlet adatai azt mutatják, hogy ez a könnyebbnek ítélt percepció működés sok esetben még iskoláskorban sem tökéletes. A megfelelő percepció működés hiányában a gyermekek mind több kudarcot élnek át, s már nemcsak elmaradásaik, de azok pszichés következményei is hátráltatják a tanulási folyamataikat. A beszédészlelési mechanizmus életkor-specifikus működésének ellenőrzése, valamint a meglévő zavarok csökkentése a megfelelő fejlesztéssel elengedhetetlen.

Irodalom

- Bishop, Dorothy V. 1992. The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 33. 3–66.
- Csépe Valéria 2003. EN, a sokat ígérő negativitás. Eseményhez kötött agyi potenciálok a beszédészlelés és a diszlexia vizsgálatában. *Magyar Pszichológiai Szemle LVIII.* 243–265.
- Evans, Julia L. 2001. An emergent account of language impairments in children with SLI: implications for assessment and intervention. *Journal of Communication Disorders* 34. 39–54.
- Gósy Mária 2000. *A hallástól a tanuláshoz*. Nikol Kkt., Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2006. *GMP-diagnosztika*. Nikol Kkt., Budapest.
- Grunwell, Pamela 1987. *Clinical phonology*. Croom Helm, London.
- Hörmann, Hans 1971. *Psycholinguistics*. Springer, Berlin–Heidelberg–New York.

- Ingram, David 2001. Toward a theory of phonological development. In Kreidler, W. (ed.): *Phonology*. Routledge, London, 60–79.
- Smith, Bruce L. – Kenney, Mary Kay 1998. An assessment of several acoustic parameters in children's production development: longitudinal data. *Journal of Phonetics* 26. 95–108.
- Winitz, Harris – Bellerose, Betty 1968. Relation between sound discrimination and sound learning. *Journal of Communication Disorders* 1. 215–235.

A kutatás az OTKA T 049426 számú támogatás keretében folyt.

AZ OLVASÁSZAVAR ÉS A BESZÉDFELDOLGOZÁSI FOLYAMATOK ÖSSZEFÜGGÉSEI

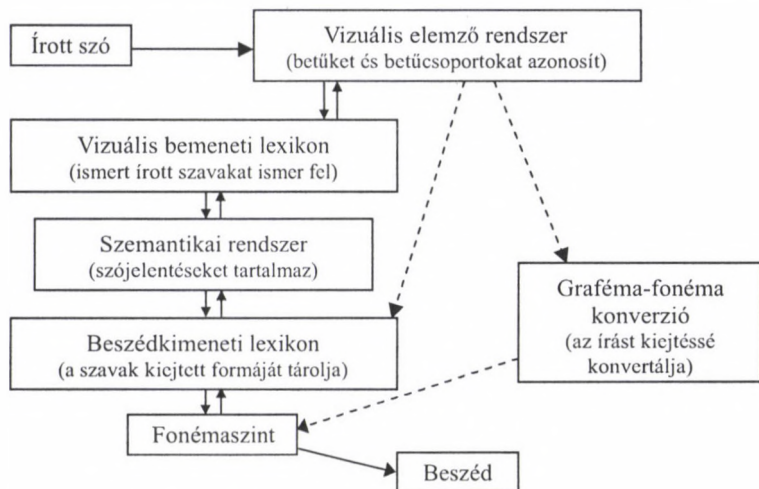
Imre Angéla

Bevezetés

Az olvasás olyan dekódolási készség, amelynek során a leírt szavakat beszélt (kiejtett) szavakká alakítjuk át (Perfetti 1986). Ez a folyamat, természetesen ép intelligenciát, hallást és látást feltételezve, igen szoros kapcsolatban van az anyanyelvvel: az olvasás tanulásának feltétele a megfelelő szintű anyanyelvi tudás, ezen belül is elsősorban a beszédészlelési és beszédmegértési folyamatoknak van döntő szerepe. Az olvasás jelenti egyfelől az észlelést és a megértést, így kimondható, hogy sikerességét a beszédpercepció folyamatai alapvetően meghatározzák. A jól olvasás azért is nagyon fontos, mert az olvasás technikai tudása és az olvasottak értése mindenképpen a szóbeli és az írásbeli kommunikációt szolgálja, valamint elengedhetetlen a sikeres tanuláshoz. Napjainkban számos gyermeknek kell részt vennie logopédiai, illetve fejlesztő pedagógiai foglalkozáson, mert a tanulásban nem eredményesek: nem értik a szöveges feladatokat, nem ismerik fel az összefüggéseket, illetve nem képesek kiemelni az olvasottakból, elhangzottakból a lényegét. A jelenség háttérében általában olvasásértési vagy olvasástechnikai zavar áll. Egy, a 80-as években elvégzett felmérés szerint a 4. osztályosok 22%-a nem érte el az értő olvasásban a 40%-os teljesítményt (Kádárné Fülöp 1985); a Monitor-vizsgálatok adatai alapján a felső tagozatban még visszaesés is jelentkezhet (Horváth 1997). Ennek tapasztalataink szerint leggyakrabban beszédfeldolgozási zavar az oka.

Milyen mértékben befolyásolja a beszédészlelési és beszédmegértési részfolyamatok működése az olvasási folyamatot? A probléma megértéséhez először tekintsük át, hogyan is olvasunk. A szakirodalomban ismeretes számos modell közül az 1. ábrán Ellis és Young (1988) olvasásmodellje látható, amely három lehetséges olvasási utat feltételez. A gyakorlott olvasókra a folytonos nyilakkal jelzett út jellemző, vagyis az írott szó betűit, betűcsoportjait a vizuális elemző rendszer azonosítja, majd a vizuális bemeneti lexikon felismeri az ismert írott szavakat, ezután következik a szemantikai rendszer, amely a szójelentéseket tartalmazza. Ezután aktiválódik a beszédkimeneti lexikon, amely a szavak kiejtett formáját tárolja, majd a fonémaszint. Hangos olvasás esetén pedig a végső lépés a meghangosítás. Fontos megemlíteni, hogy ha elmarad a vizuális bemeneti lexikon aktiválása, és a szavak jelentését nem ismeri vagy nem dolgozza fel a szemantikai rendszer, akkor nem beszélhetünk az olvasottak megértéséről (Eysenck–Kean 2003). Ehri (1994) feltételezi, hogy a gyakorlottsággal kialakul először az írott

kép és a hangzó alak egysége, majd ezen egység és a jelentés egysége, amely azonnal aktiválódik értő olvasáskor.



1. ábra

Ellis és Young olvasásmodellje

(Ellis és Young 1988 alapján Eysenck–Kean 2005: 331)

A hallott és az írott nyelv dekódolásához elengedhetetlen feltétel az emlékezet, a hallás (írott szöveg esetében a látás) és az észlelési folyamatok tökéletes működése. A szöveg tartalmának megértéséhez képesnek kell lennünk a következtetések levonására, a szövegszerkezet felismerésére és a saját megértési folyamataink ellenőrzésére. Bloomfield (1942) óta tudjuk, hogy az olvasás elsajátításában a nyelvi folyamatoknak meghatározó jelentőségük van. A releváns szakirodalom többsége a nyelvi tudatosság szerepét hangsúlyozza, amelynek egyes szintjei párhuzamba állíthatók a beszédmegértés hierarchikus felépítésű modelljének megfelelő szintjeivel. Ezek közül a legfontosabb a fonématudat (angol szakkifejezéssel: *phonemic awareness*), amely „a szavakat alkotó hangok (fonémák) tudatát jelenti” (Harris–Hodges 1995, idézi A. Jászó 2006: 151); a fonémareprezentációnak az olvasás kezdetekor már stabilnak kell lennie (Csépe 2006). A Jászó saját meghatározása szerint „a nyelvi tudatosság olyan intuitív analízis képesség, melynek segítségével megfigyeljük beszédünket, s alkotórészeire bontjuk. [...] Előfeltétele az olvasás-írás megtanulásának, de az olvasás és írás tanulása közben tovább fejlődik, és magasabb szintre kerül. Tökéletes minősége tehát következménye az olvasás és az írás megtanulásának” (2006: 153).

A beszédfeldolgozási zavarokon kívül a nem megfelelő olvasástanítási eljárások is felelősek lehetnek az olvasási gyengeségért, bár nehéz meghatározni, mi-

lyen arányban (Subosits 2001). A jó olvasás elsajátítását az alapozhatja meg, ha a valódi olvasástanulást megelőzi egy analízáló-szintetizáló előkészítő szakasz, azaz a gyermekek felbontják a mondatokat szavakra, a szavakat szótagokra, majd hangokra. Az 1978-ban bevezetett olvasástanítási módszerek azonban a nyelvi tudatosságnak a kialakulását nem segítik elő, hiszen szinte eltűnik a hangoztatás: elmarad a klasszikus előkészítési szakasz, a gyermekeket némán olvasztatják, a betűtanítást követi a hangtanítás, és nem fordítva, valamint nincs szótagolás sem. Csépe és munkatársai kísérletet terveztek annak megállapítására, számít-e, hogy milyen módszerrel tanulunk meg olvasni, azaz a fonológiai tudatosság stabilitása kialakul-e akkor is, ha ezt az alkalmazott tanítási módszer nem támogatja. A kísérletben két 1. osztály kisiskolárai vettek részt. A két osztályban eltérő olvasástanítási módszert alkalmaztak a pedagógusok: az egyikben a szóképfolvasást, azaz a globális előprogrammal kezdő módszert, míg a másikban a szótagoló, elemző-hangoztató módszert. A gyermekeket aszerint osztották alcsoportokba, hogy milyen teljesítményt nyújtottak a beszédhang-differenciálási próbában. A követéssel vizsgált eredményei rámutattak, hogy míg az elemző-hangoztató módszerrel tanuló osztályban a jó diszkriminációs képességekkel rendelkező gyermekeknél a két év alatt határozott a fejlődés, és a rosszul diszkriminálónál felzárkózás mutatkozik, addig a globális módszerrel tanuló csoportban alig lehetett teljesítményváltozást megfigyelni (Csépe 2006). Más vizsgálatok megállapították, hogy azok között, akik a szóképes módszerrel tanultak olvasni, nagyobb arányban találhatunk gyenge helyesírókat is (A. Jászó 2003). Az 1989 utáni olvasástanítási programoknál újra figyelembe veszik a hangoztatás-percepció szempontjait, visszakerül a szótagoltatás, előtérbe kerül a szövegértést segítő szemléltetés, de a beszédre alapozott tanítás nem valósul meg olyan mértékben, mint az 1978 előtt történt (A. Jászó 2001 és 2006).

Ha az olvasás minősége nem megfelelő, az olvasási technika felülvizsgálata mellett is elsősorban a beszédészlelést kell megvizsgálni, és annak irányából érdemes megtervezni a fejlesztést is. Hogyan történik a beszédfeldolgozás? Röviden ez így írható le: A beszédképzés következményeként kialakuló hanghullámok, akusztikus rezgések eljutnak a hallgatóhoz, a hallószerv, azaz a fül felfogja ezeket a rezgéseket, és elkezdődik a hallási feldolgozás, majd a kódolt információk továbbítódnak az agyba, ahol egy bizonyos területen az előbb sematikusán vázolt folyamat végén megtörténik a megértés. Elmondható tehát, hogy a beszédfeldolgozás tökéletes működéséhez elengedhetetlen a fülünk és az agyunk épsége, valamint a két szervben folyó részfolyamatok összehangolt munkája. Ha a beszédfeldolgozási részfolyamatok a működésben elmaradnak vagy zavart mutatnak, ez megnehezíti az olvasás- és írástanulást, s ennek negatív hatása lehet a többi tanulási folyamatra is (Gósy 2005).

Anyag, módszer, kísérleti személyek

Jelen kutatásunkban arra a kérdésre kerestük a választ, hogyan működnek a beszédfeldolgozás egyes szintjei azoknál a gyermekeknél, akik olvasási nehéz-

séggel küzdenek. A gyermekek beszédészlelési és beszédmegértési folyamatának objektív felméréséhez olyan teszteljárást kerestünk, amelynek feladatai a lehető legszélesebben lefedik a hierarchikus beszédfeldolgozási folyamat részfolyamatait. A Gósy Mária által kidolgozott GMP percepciósi teszt (Gósy 2006b) minden részfolyamat működését vizsgálja, Magyarországon a beszédfeldolgozás vizsgálatára ez a legtöbbször használt mérőeljárás. A tesztsorozatban elért eredmények ismeretében megítélhetjük, hogy a gyermekek beszédészlelése megfelelő-e az életkoruknak vagy nem, hol jelentkezik a beszédészlelési mechanizmusban elmaradás vagy zavar, milyen mértékű és kiterjedésű az elmaradás vagy zavar, milyen kompenzáló stratégiákat alkalmaznak a gyermekek, illetve melyek a korrigálás lehetőségei.

Kutatásunkban azt is megvizsgáltuk, hogy az egyes részfolyamatok mennyire képesek önállóan működni, vagy felfedhető-e működésükben valamilyen összefüggés; ehhez az említett GMP-diagnosztika tesztjei közül a beszédészlelést vizsgáló GMP2, GMP3, GMP4, GMP5 és GMP10, a beszédhang-differenciálást mérő GMP17, a transzformációs észlelést vizsgáló GMP18, valamint a mondat-és szövegértést ellenőrző GMP16 és GMP12 alteszteket alkalmaztuk.

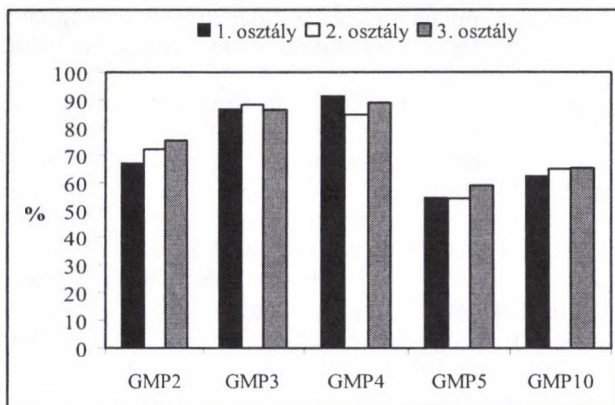
A vizsgálatban 1–3. osztályos diákok vettek részt, évfolyamonként 50-50 fő, összesen 150 tanuló. A gyermekek hagyományos oktatási rendszerű iskolába járnak; logopédus, illetve fejlesztő pedagógus olvasási nehézséggel küzdőknek diagnosztizálta őket, de nem állapítottak meg náluk diszlexiát. Feltételezéseink szerint az általunk választott teszteljárással kimutatható, hogy a kísérletben részt vevő gyermekeknél az olvasási nehézség hátterében minden esetben beszédfeldolgozási zavar is húzódik. A kapott eredményeket statisztikailag is elemeztük, a Pearson-próbához az SPSS 13.0 verzióját használtuk.

Eredmények

Az akusztikai, fonetikai és fonológiai szint részfolyamatainak megfelelő működése teszi lehetővé a jó beszédészlelést. Az ép anyanyelv-elsajátítási működésnek megfelelően a szóazonosítás (GMP3) jobb a mondatazonosításnál (GMP2), de a mért teljesítmények azt mutatják, hogy a vizsgált csoportok mindegyike mindenhol elmarad az életkorában elvárható teljesítménytől (a GMP2-ben és a GMP5-ben a 6 éveseknél 90%, az ennél idősebeknél, illetve a többi altesztben minden korcsoportnál 100% az elvárt szint). Fejlődés az életkor előrehaladtával a GMP2-ben állapítható meg, de a többi beszédészlelési folyamatnál visszaesést vagy stagnálást tapasztaltunk. A legnagyobb elmaradás a fonológiai szintnél (GMP5) látható, itt még a 3. osztályosok is átlagban 60% alatt teljesítenek (az átlagteljesítmény az 1. osztályban 55, a 2. osztályban 54, míg a 3. osztályban 59%-os) (2. ábra).

A beszédfolyamatban csak akkor tudunk sikeresen részt venni, ha megértjük és betartjuk a nyelvnek azon szabályát, hogy a beszédhangok sorozatának sorrendisége mindig meghatározott, csak ez teszi lehetővé, hogy mind a beszélő, mind a hallgató mindig egyértelműen azonosítani tudja a konvencionális jelsort

(Gósy 1996). A szeriális észlelés (GMP10) ép működése biztosítja, hogy a gyermekek képesek legyenek az elhangzó beszédhangok, hangkapcsolatok felismerésére és reprodukálására, illetve a jelentéssel nem bíró értelmetlen hangsorok pontos ismétlésére. Ez utóbbi szükséges például az idegen nyelvek tanulásához is. A mért teljesítmények azt mutatják, hogy a sorozatészlelés szintje minden korcsoportban gyenge, az elmaradás mértéke jelentős, 37, illetve 35 százalékpontnyi.

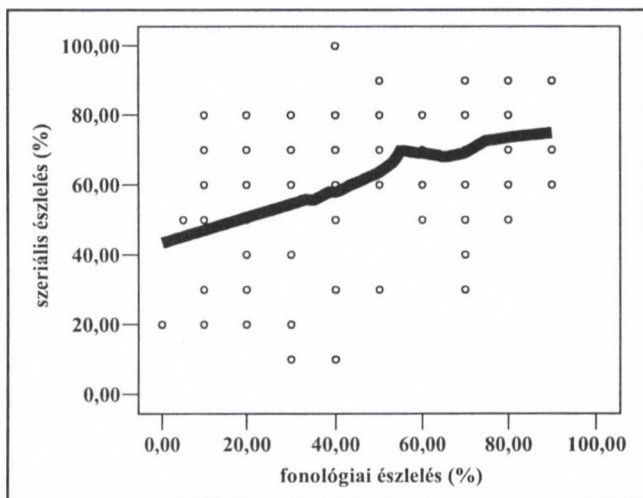


2. ábra
Beszédészlelési teljesítmények

A Pearson-próba 0,01 szinten szignifikáns és közepesen erős korrelációt mutat ($r = 0,460$) a fonológiai és a szeriális észlelés között (3. ábra), ami azt jelenti, hogy ezek a szintek még nem működnek jól önállóan, az egyik részfolyamat függ a másik sikerességétől. Ez pedig nem jó, hiszen 7 éves kortól már a két részfolyamat szétválását várjuk el, amint azt az újabb kutatások igazolták (Gósy 2006a).

A beszédpercepció egyik fontos mérföldköve, amikor a gyermekek képesek lesznek az anyanyelvükben létező egyes hangokat egymástól függetlenül is észlelni, illetve azokat egymással összevetni. A beszédhangok differenciálásának képessége elengedhetetlen ahhoz, hogy a megfelelő szókincs kialakuljon és továbbfejlődjön, de fontos és kritikus előfeltétele az olvasás és írás elsajátításának is (Fazekasné Fenyvesi 2000). 6 éves korra elvárható, hogy a gyermekek az anyanyelvük beszédhangjait meg tudják különböztetni egymástól, függetlenül attól, hogy két hangzó milyen mértékben és milyen akusztikai tulajdonságokban tér el egymástól. Kivételt csak a magán- és mássalhangzók időtartambeli különbségének pontos felismerése jelenti, de 7 éves korra már ez sem okozhat gondot. Ha a beszédhangok megkülönböztetésének képessége nem válik stabilá legkésőbb 7 éves korra, ez akadályozza az anyanyelvi szókincs bővülését, nehe-

zíti az olvasás, írás, helyesírás elsajátítását, de nehézséget okoz más nyelvek tanulásában is.



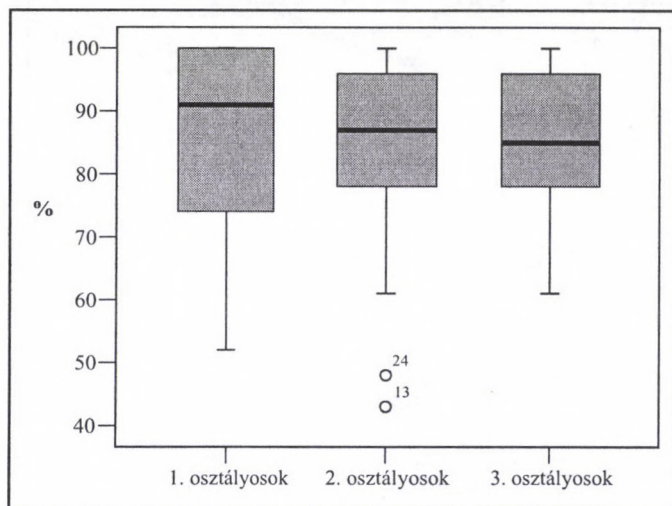
3. ábra

A fonológiai (GMP5) és a szeriális észlelés (GMP10) összefüggése

A beszédhang-differenciálási altesztben (GMP17) gyakorlatilag nincs különbség az egyes évfolyamok teljesítményei között: a gyermekek 83–85%-os átlagteljesítményt nyújtottak. A 4. ábrán szemléltettük az adatok szóródását évfolyamonkénti (így láthatók a leggyengébb és legjobb teljesítményt jelentő százalékos értékek), valamint az egész évfolyam mediánját.

Ha a transzformációs észlelés megfelelően működik, akkor a gyermekek képesek lesznek arra, hogy bizonyos akusztikus információkat átvigyenek valamilyen vizuálisan észlelhető jelenségre, és a későbbiekben az egyiket azonosítani tudják a másik alapján. Ez a folyamat az alapja annak, hogy a beszédhangokat az őket szimbolizáló betűknek megfeleltessük. A tökéletes működés négy feltétele: 1. beszédhangnyi időtartamú (kb. 50–200 ms) hangjelenségek azonosítása, 2. vizuális jelenségek felismerése, 3. a feltételes reflex alapján kialakított kapcsolat megértése és fenntartása, valamint 4. mindezek ismételt működtetése (Gósy 1996: 44). Az alteszt négy próbája a transzformációs észlelés különböző fokozatainak meglétét vizsgálja. Az elsőnél egy jelet (beszédhangot) kell megfeleltetni egy alaknak (kockának), de a gyermekeknek szükségük van a hibátlan megoldáshoz bizonyos ritmusészlelésre is. A második próba a tulajdonképpeni transzformációs percepció: egy új jelnek a már ismert alakon kívüli másik alakot kell megfeleltetni. A harmadik próba sikeres megoldásában részben a hangzás időrendisége meghatározó (szerialitás), részben a „térbeli irányra vonatkozó

szeriális észlelés”. A negyedik próba pedig csak akkor oldható meg tökéletesen, ha a gyermekek a jó transzformációs és szeriális észlelésen túl hibátlanul azonosítják a beszédhangokat, és jó a rövid távú verbális memóriájuk. Ha a gyermekek transzformációs észlelési szintje az adott életkorra nem vagy nem megfelelő szinten fejlődik ki, akkor nehezebben feleltetik meg a beszédhangokat a betűknek, ami betűfelismerési zavarokhoz, illetve olvasási nehézséghez, írászavarokhoz vezethet.



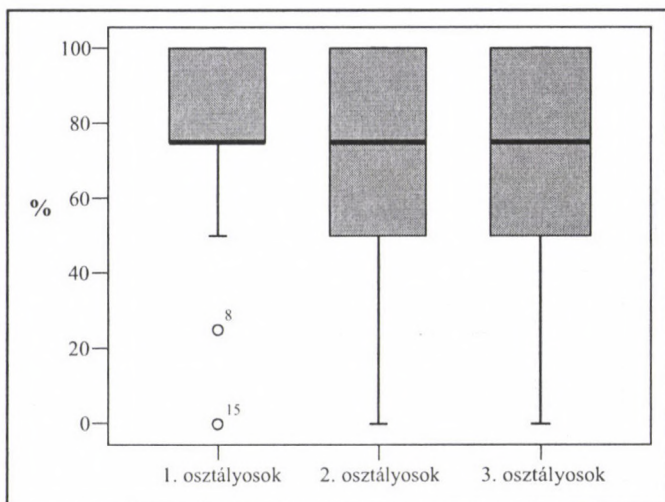
4. ábra

Beszédhang-differenciálási teljesítmények (szóródás és medián)

Ép fejlődés esetén hatéves korban 75%-ot, hatéves kortól 100%-ot várunk el. Az általunk mért három évfolyam 77, 71 és 70%-os átlagteljesítményt nyújtott. Az 5. ábráról kiderül, hogy minden évfolyamban volt diák, aki egyik résztesztet sem tudta megoldani, de mindenhol volt hibátlan teljesítmény is. Az adatok jól mutatják, hogy a transzformációs észlelésnél sem lehet az életkor növekedésével javuló teljesítményt megállapítani. Sőt a harmadikosok átlageredménye az elsőskéhez képest még 7 százalékponttal gyengébb is lett. Feltételezhetően ezek az eredmények jelentkeznak a gyenge olvasási és nem megfelelő írási teljesítményekben.

A beszédfeldolgozás során a beszédészlelésre épül a beszédmegértés szintje. Megfelelő működését számos tényező befolyásolhatja; de főleg kisebbeknél meghatározó szempont lehet, hogy a gyermekeknek milyen hosszúságú szöveget kell feldolgozniuk. A mondat- és szövegértést érdemes egymástól függetlenül is vizsgálnunk, hiszen az életkornak megfelelő szövegfeldolgozás nem feltétlenül jelenti a mondatértés jó működését, és ez fordítva is igaz, azaz a mondatértés tel-

jesítménye nem lesz minden gyermek esetében azonos a szövegértés teljesítményével.



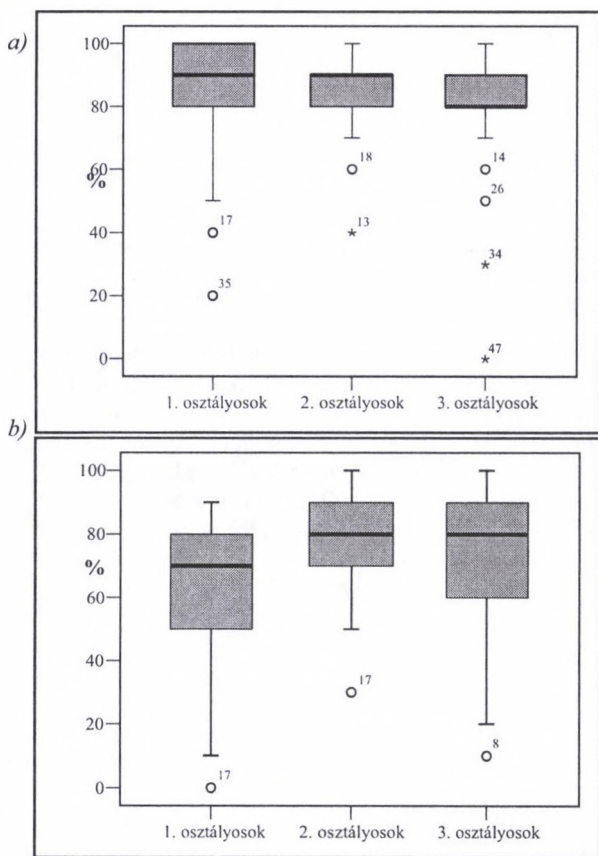
5. ábra

Transzformációs észlelési teljesítmények (szóródás és medián)

A mondatértési nehézségek elsődleges oka az, hogy a gyermekek nem megfelelően sajátították el az anyanyelvükre jellemző grammatikai és szintaktikai struktúrákat, illetve szemantikai egységeket. Ha a gyenge teljesítményeik nem javulnak, az ahhoz vezet, hogy a szabályokat, definíciókat és szöveges feladatokat, azaz a rövid nyelvi információkat nem fogják tudni megérteni, és ez gátja a tanulási folyamatoknak. A szövegen belüli mondatértési zavar megnehezítheti az egész szöveg megértését is. Az adott nyelvi jelek helyes feldolgozásán kívül elengedhetetlen az ismeretek beépítésének és alkalmazásának képessége, az ok-okozati összefüggések, az időviszonyok helyes felismerése és jó logika ahhoz, hogy a gyermekek az életkoruknak megfelelő szövegértéssel rendelkezzenek (Gósy 2006b). Ismert tény, hogy amíg a mondat nagyságrendű közlések megértése megtörténhet az asszociációs szint működése nélkül, addig a szövegértésbe már az említett szintnek is be kell kapcsolódnia: sémákat, szerkezeteket és különböző forgatókönyveket (idegen szóval *script*-eket) alkalmazunk, hogy egy történetet minél jobban megérthessünk, s adott esetben később visszaidézhesük (Eysenck–Keane 2003).

Ha a szövegértés folyamata nem az életkornak megfelelő szinten működik, akkor sérül a gyermekek beszédprodukciója, olvasásértési és tanulási problémák léphetnek fel, de adva van a gyermekek kognitív szintjének és a pszichés sérülésének lehetősége is.

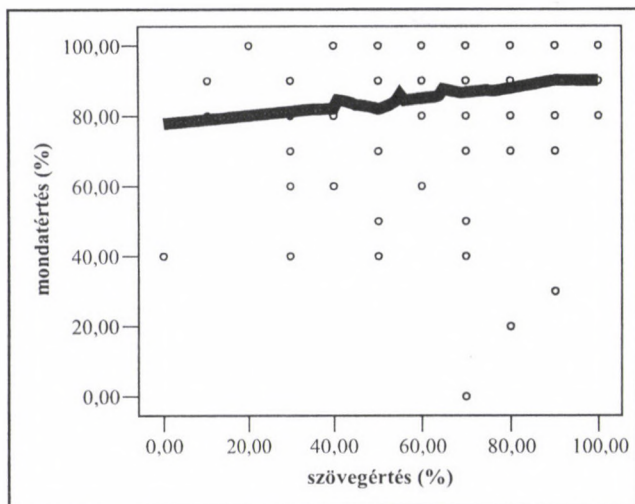
A mondat- és szövegértés mechanizmusa független lehet egymástól. Ezzel együtt érdemes összevetnünk a gyermekek mondat- és szövegértési teljesítményét és meghatározni a részfolyamat kapcsolatát. Az anyanyelv-elsajátítás kezdetében általában a mondatértés magasabb szintű, mint a szövegértés, de később a két folyamat között már nem lesz különbség, ha a gyermekek beszédfeldolgozása megfelelően fejlődik (Gósy 2000: 79). Az általunk vizsgált 1. osztályosok mondatértése átlagosan 84%-os, a másodikosoké 87, míg a harmadikosoké 82%-os volt. A szövegértés eredményei évfolyamonként átlagosan: 64, 75, illetve 70%. Látható, hogy mindenhol a mondatértés volt sikeresebb (a különbség 23, 7 és 12 százalékpontnyi), a 8 éves kor körül egészséges fejlődés esetén elvárt kiegyenlítődés még nem következett be (6. ábra).



6. ábra

a) Mondat- és b) szövegértési teljesítmények (szóródás és medián)

A hallás utáni mondat- és szövegértés összefüggésének vizsgálatához elvégeztük a Pearson-próbát, amely 0,05 szinten szignifikanciát mutat, ugyanakkor korreláció a két alteszt teljesítményei között alig mutatható ki ($r = 0,189$). Ez az adat arra utal, hogy a mondatértés fejlettsége nem függ a szövegértés fejlettségétől, ami a fejlődésben jelentős előrelépés (7. ábra).



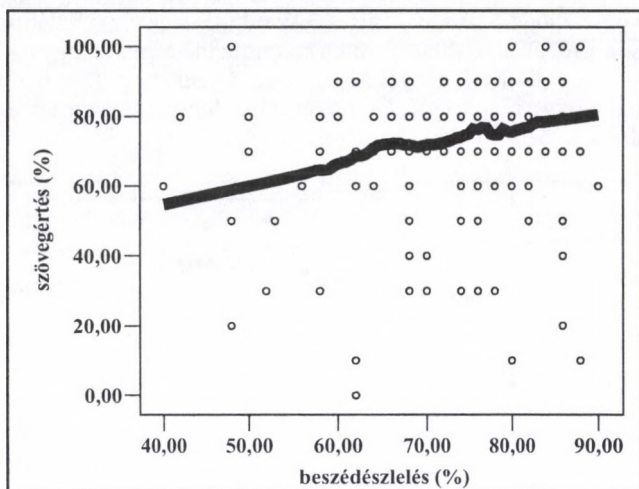
7. ábra

A hallás utáni mondat- és szövegértés összefüggése

A beszédészlelés és a hallás utáni szövegértés nagyon gyenge korrelációja (0,01 szinten $r = 0,224$) azt jelenti, hogy az egész vizsgálati csoportot tekintve csak kismértékben függ a hallás utáni szövegértés az észleléstől (8. ábra). Ez az eredmény is jobb, mint a fonológiai és a szeriális észlelés között megállapítható összefüggés, de gyengébb az előbb bemutatott eredménytől.

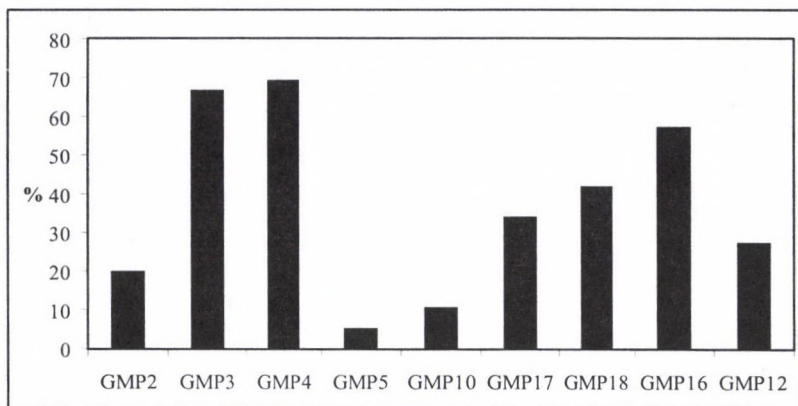
Összefoglalás

A számadatok igazolják hipotézisünket, miszerint az olvasási nehézséggel küzdő gyermekeknél minden esetben találunk beszédfeldolgozási zavarra utaló jeleket. A vizsgált olvasási nehézséggel küzdő gyermekek akusztikai, fonetikai és fonológiai szintjének működése jelentősen elmarad az adott életkorban elvárt szinttől: a 150 gyermek között nem volt senki, akinél mindegyik részfolyamat az életkorára jellemző elvárt szinten működött volna. A 9. ábra szemlélteti, hogy a gyermekek hány százaléka teljesített az egyes altesztokban az elvárt szinten. A legnagyobb elmaradás a fonológiai szintet és a szeriális észlelés szintjét vizsgáló altesztben volt.



8. ábra

A beszédészlelés és a hallás utáni szövegértés összefüggése



9. ábra

Az elvárt szinten teljesítők aránya

A korrelációs vizsgálatok rámutatnak, hogy az egyes szintek még nem tudnak megfelelően egymástól függetlenül működni. A beszédhang-differenciálási és a transzformációs észlelési alteszt eredményei azt jelzik, hogy az életkor növekedésével sem mutatható ki a vizsgált csoportoknál fejlődés, holott ezt várnánk az intézményes oktatásban való részvételtől. Az olvasási zavarral küzdő gyermekek esetében a jelenleg alkalmazott olvasástanítási módszerek feltehetően nem alkalmasak ennek a célcsoportnak, s feltétlenül szükség van célzott fejlesztésre.

A szövegértés fejlődését segíti a gyakori, jól megválasztott szövegen történő gyakorlás, erre érdemes nagyobb hangsúlyt fektetni az olvasásórákon, hiszen a szövegértés fejlődése elősegíti más tantárgyak sikeresebb elsajátítását is a feladatértés miatt. A nem megfelelően működő részfolyamatok felismeréséhez és az eredményes fejlesztéshez pedig elengedhetetlen a tanítók és a fejlesztő pedagógusok további közös együttműködése.

Irodalom

- Bloomfield, Leonard 1942. Linguistics and reading. *The Elementary English Review* XIX. 125–130, 183–186.
- Csépe Valéria 2006. *Az olvasó agy*. Akadémiai Kiadó, Budapest.
- Ehri, Linnea C. 1994. Development of the ability to read words, Update. In Ruddell, Robert B. – Ruddell, Martha R. – Singer, Harry (eds.): *Theoretical models and processes of reading*. 4th edition. International Reading Association, Newark, Delaware, 323–358.
- Ellis, Andrew W. – Young, Andrew W. 1988. *Human cognitive neuropsychology*. Lawrence Erlbaum Associates Ltd., London.
- Eysenck, Michael W. – Keane, Mark T. 2003. *Kognitív pszichológia*. Nemzeti Tankönyvkiadó, Budapest.
- Fazekasné Fenyvesi Margit 2000. A beszédhanghallás kritériumorientált fejlesztése. <http://www.oki.hu/oldal.php?tipus=cikk&kod=2000-07-km-fazekasne-beszedhanghallas>. A letöltés ideje: 2005. december 10.
- Gósy Mária 1996. A transzformációs észlelés fejlődése és zavarai. *Beszéd kutatás* 1996. 44–57.
- Gósy Mária 2000. *A hallástól a tanulásig*. Nikol Kkt., Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Gósy Mária 2006a. *A beszédpercepció zavarok okainak és sajátosságainak összefüggése*. Előadás. IX. Pszicholingvisztikai Nyári Egyetem, Balatonalmádi.
- Gósy Mária 2006b. *GMP-diagnosztika*. Nikol Kkt., Budapest.
- Harris, Theodore L. – Hodges, Richard E. (eds.) 1995. *The literacy dictionary. The vocabulary of reading and writing*. International Reading Association, Newark, Delaware.
- Horváth Zsuzsanna 1997. Szövegek és olvasóik. Helyzetkép a tanulók szövegértéséről. In Vári Péter (szerk.): *Monitor '95. A tanulók tudásának felmérése*. Országos Közoktatási Intézet, Budapest, 107–201.
- A. Jászó Anna (szerk.) 2001. *A magyar olvasás története*. Osiris Kiadó, Budapest.
- A. Jászó Anna 2003. *Csak az ember olvas*. Tinta Könyvkiadó, Budapest.
- A. Jászó Anna 2006. *Az olvasás múltja és jelene*. Trezor Kiadó, Budapest.
- Kádárné Fülöp Judit 1985. *Olvasás és kommunikáció*. OPI, Budapest.
- Perfetti, Charles A. 1986. Cognitive and linguistic components of reading ability. In Foorman, Barbara R. – Siegel, Alexander W. (eds.): *Acquisitions of reading skills*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 11–41.
- Subosits István 2001. *A beszéd rendellenességei*. Egyetemi Fonetikai Füzetek 30.

A FÁRADÓ BESZÉD MINT BETEGSÉGTÜNET

Horváth M. Szabolcs – Magony Edit – Gósy Mária

Bevezetés

Az interdiszciplináris kutatások leggyakrabban akkor és úgy jönnek létre, ha a kutatás tárgya több tudományterületet is érint, avagy ha a felvetett kérdések megválaszolása nem oldható meg egyetlen tudományon belül. A beszéd jellegzetesen több tudományterülettel kapcsolatos; elsősorban a fonetikával és a pszicholingvisztikával (Gósy 2004, 2005); de a neurolingvisztika, a neuropszichológia, sőt bizonyos részeivel a beszédtechnológia, avagy a fül-orr-gégészet foglalkozik. Az ép beszéd folyamatok pontosabb megismerését segíti a normálistól eltérő működések tanulmányozása és leírása. A patológiás változások törvényszerűségei és tendenciái valamiféle belső kapcsolatot tartanak a normális funkciókkal; az eltérés a normális kognitív struktúra egyfajta felbomlása, amelynek elemzésével következtetések fogalmazhatók meg az ép működésekre vonatkozóan. A gyakorlat szempontjából különösen fontos a patológiás beszéd azonosítása és megfelelő elemzése, mert ennek révén egy adott betegség vagy kóros folyamat diagnosztizálása lehet biztosabb. A beszéd hallható megváltozása a legkorábban jelentkező tünetek egyike lehet a legkülönbözőbb betegségekben. Az éptől eltérő beszéd folyamatok vizsgálata tehát magától értetődően interdiszciplináris megközelítést igényel, amelyben beszédkutató, fonetikus, orvos, neurológus együttműködése hozhat kiváló eredményt.

A beszédzavarok mindig együtt járnak a beszélő/hallgató pszichés problémájával, hiszen a beszélő/hallgató folyamatosan szembesül saját kommunikációs nehézségeivel. Érinthetnek egy-egy részfolyamatot, de kiterjedhetnek a teljes beszédre is; mértékük a legkülönbözőbb lehet. A beszédzavarok felismerését és különösen egy adott betegséghez kapcsolását nagymértékben megnehezíti a beszélő/hallgató relatív toleranciája, valamint az, hogy nemegyszer keverednek az organikus és a funkcionális, illetőleg a pszichés problémák. Némely beszédzavar feltűnőbb lehet a környezetnek (a hallgatóknak), mint magának a beszélőnek; a beszélők pedig sok esetben a fáradtságukra, kimerültségükre vezetik vissza a verbális kommunikáció során jelentkező zavarait. A nyelvi és a beszédzavarok tüneteinek a legkülönbözőbb kombinációkban és a legkülönbözőbb betegségekhez, szindrómákhoz kapcsolódóan jelennek meg. Lehetnek kísérő jelenségei az alapbetegségnek, mint például a kiejtési nehézség a Parkinson-kórban vagy a grammatikai megformálás nehézsége demenciában. Lehetnek azonban a kórképek előjelző jelenségek, amilyen például a „kifáradó” artikuláció a myasthenia gravisban (= MG). Ez a betegség a neuromuszkuláris junció működési zavara, sú-

lyos izomgyengeség, az izomműködés szokatlanul gyors kimerülése, amit az ingerületátvivő anyag mennyiségének különböző okú csökkenése okoz. Tekintettel arra, hogy a betegség egyértelmű felismerése meglehetősen nehéz, ezért minden adalék fontos lehet, amely időben ráirányíthatja a figyelmet. Közülük elsőrendű szerepe van a beszéd megváltozásának. Ezt különösen az első esetismertetésünk igazolja.

A súlyos izomgyengeség (myasthenia gravis) története

A betegséget – amelyet hullámzó lefolyású izomgyengeség és az izmok kóros fáradékonysága jellemez – Willis írta le elsőként 1672-ben. A beteg izomműködés-teljesítményét úgy jellemezte, hogy az „néma, mint a hal”. 1879-ben Erb úgy találta, hogy a myasthenia gravis más típusú izomgyengeséget okoz, mint az az egyéb betegségekben megszokott. 1893-ban Goldflam kiegészítette Erb leírását a tüneti képről, s néhány évvel később Jolly használta elsőként a *myasthenia gravis pseudoparalytica* megjelölést. 1901-ben Weigert összefüggést talált a csecsemőmirigy (thymus) túltengése és a betegség kialakulása között, míg 1913-ban Schumacher és Roth elvégezte az első csecsemőmirigy-eltávolítást. 1934-ben Mary Walker a betegség gyógyítására physostigmint alkalmazott, s ez az acetylcholinesterase-gátló készítmény jelentős izomerősödést eredményezett a betegnél.

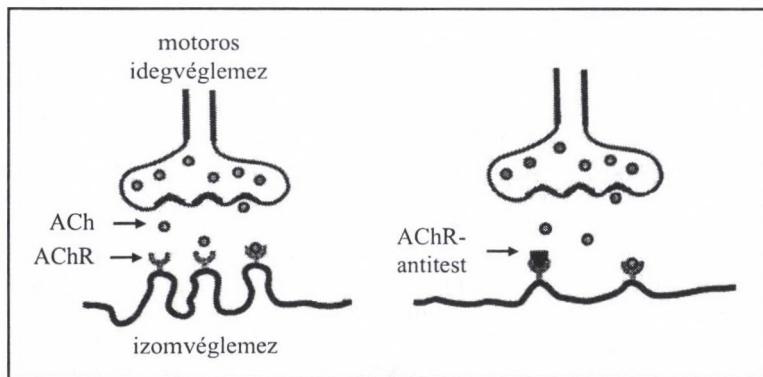
Harvey és Masland (1941-ben) közölte az első elektrofiziológiai leletet, amely szerint az ideg ismétlődő ingerlésekor izomfáradás figyelhető meg, ami az amplitúdó csökkenésével kifejezhető. Simpson 1960-ban leírta, hogy a MG kialakulása immunológiaiilag közvetített anyag támadása az izomvéglemezen. 1971-ben Patrick és Lindstrom állatkísérletben megfigyelte a betegség hullámzó lefolyását, s megállapította, hogy a változékonyságot mutató neostigmin-érzékeny izomgyengeség és az acetylcholinreceptor (AChR) fehérje immunizálása között kapcsolat van. 1976-ban Lindstrom és munkatársai megtalálták a szérumban az AChR elleni antitestet, és meghatározták a posztzinaptikus izommembránon az immunkomplex szerepét, amely a patognosztikában jelentős előrelépést hozott. Eszerint az immunkomplex kötődése az ACh-receptoron indítja el a kóros élettani folyamatokat. Mivel az ACh nem tud kötődni a receptorhoz, így nem jön létre ingerületátvitel, s az izom nem kapja meg az összehúzóadáshoz szükséges információt. Az izomzat bizonyos izomcsoportjai nem húzódnak össze, azaz fáradékonyabb lesz (vö. Szobor 1983, Bradley et al. 2000).

A magyar orvoskutatók közül Albert foglalkozott e problémával; 1983-ban 30 év kutatási eredményeit foglalta össze. 675 súlyos izomgyengeségben szenvedő beteg kezelése során kidolgozta ennek a betegségnek a diagnosztikáját, hosszú távú gondozását és magyarországi epidemiológiáját (Szobor 1977, 1983). Az új epidemiológiai adatok szerint 100 000 ember közül megközelítőleg 14 szenved súlyos izomgyengeségben. (A magyarországi incidenciasetek száma kb. 1000. Az USA adatai szerint ott ez a szám mintegy 36 000.) A betegség azonban aluldiagnosztizált, ami azt jelenti, hogy valószínűleg sajnos több a nem

felismert beteg. A nők aránya valamivel magasabb (kb. 3 : 2), a betegség többnyire az emberek húszas, harmincas éveiben kezdődik.

A súlyos izomgyengeség (myasthenia gravis) kórélettana

Az ideg-izom átmenet (neuromuszkuláris szinapszis) motoros végződésén kis hólyagsákból (vesicula) szabadul fel az ingerületet közvetítő acetilcholin, egy enzim, az acetilcholinesterase hatására, majd kis csomagokban (quantumok) megjelenik az ideg és izomvéglemmez közötti résben, végül kötődik az izommembránhoz. Ingerlés hatására sok acetilcholinquantum szabadul fel, amely depolarizálja az izomvéglemmez-régiót és az izommembránt, s ez a folyamat az izom összehúzódásához vezet. Súlyos izomgyengeségben a posztzinaptikus izommembrán felszíne egyszerűbb, kevésbé redőzött szerkezetű (1. ábra), az AChR-k száma pedig lecsökken. Noha a hólyagsákból normálisan szabadul ki az ACh, a membrán – amelyhez antitestek kötődnek – kevésbé érzékeny az ACh további felhasználására, s feltehetően valamilyen idegi impulzus hatására az izom akciós potenciálja is csökken.



1. ábra

Ideg-izom motoros végkészülék: izomvéglemmez normál működés esetén (bal oldalon) és súlyos izomgyengeségben (MG) (jobb oldalon)
(AChR: acetilcholin receptor + antitest = immunkomplex)

Az, hogy nagy valószínűséggel egy immunközvetített folyamatról van szó (Szobor 1983, Bradley 2000), bizonyítja, hogy a myasthenia gravisban gyakori más típusú autoimmun betegség is, mint a reumás ízületi gyulladás (rheumatoid arthritis), izomgyulladás (polymyositis), pajzsmirigygyulladás (thyreoiditis), bőrfarkas (lupus erythematosus). Ezt bizonyítja a betegség átmeneti újszülöttkori formája is. Hasonlóan más immunológiailag közvetített betegségekhez, súlyos izomgyengeségben szenvedő anyák újszülöttjeinek mintegy 10%-ánál átmenetileg súlyos izomgyengeség tünetei észlelhetők, amelynek oka lehet egyrészt az anyai szervezetben termelődő, de a placentán átjutó AChR ellenes antitest kötő-

dése az újszülött izomvéglemezmembránján található AChR-hoz, másrészt szintén a placentán átjutó, immunválasz csökkentésére irányuló gyógyszerek hatása (ún. immunszuppresszív gyógyszerek használatakor, beleértve a plazmacserét is). Az újszülöttkori átmeneti súlyos izomgyengeség tüneteit mutató csecsemők szérumában az AChR ellenes antitestek, az IgG és a komplement komponensek kötődése a véglemezmembránhoz mintegy 80%-ban kimutathatók speciális immunhisztokémiai módszerekkel (Szobor 1977, Bradley 2000).

A csecsemőmirigy abnormitása súlyos izomgyengeségben szenvedő betegek közel 10%-ánál mint daganatos (legtöbbször jóindulatú) elváltozás, 70%-ban pedig mint túltengéses elváltozás van jelen, az aktív immunválasz bizonyítékaként. Az ún. „germinatív centrum” a mirigyszövetnek az a területe, ahol a B-sejtek és a T-sejtek hatására antitest képződik. A csecsemőmirigyben megtalálhatók mindazok az alkotórészek, amelyek a MG kórfolyamatának kialakulásához szükségesek: az izomsejtek (myoid sejtek) – ezek az AChR antitest kialakulásához szükségesek –, az antigénsejtek és az immunkompetens T-sejtek. Nem világos még azonban, hogy azoknál a betegeknél, akiknél a csecsemőmirigy daganatos elváltozása (thymoma) észlelhető, az elektrofiziológiai eltérés általában miért súlyosabb, illetve az AChR antitest szérumszintje rendszerint miért magasabb. Hatvanéves kor fölött a thymoma jelenléte ritkább.

A súlyos izomgyengeség (myasthenia gravis) klinikai tünetei

A beteg vizsgálatának első megközelítések feltűnő az, hogy a változatos fokú izomgyengeség specifikus izomcsoportokra terjed ki. A hullámlzó izomgyengeség elsősorban a szemizmokat és a szem-garat (oculopharyngealis) izomcsoportokat érinti. Az izomerő fokát maximális erő kifejtésnél ismételtén vizsgáljuk, majd rövid szünet után a fáradtság jeleit újraértékeljük. Leggyakoribb klinikai jel egy aszimmetrikus szemhéjcsüngés, amelyet a beteg a homlokizom részvételeivel ellensúlyoz, s ez egy aggódó vagy meglepett kifejezést kölcsönöz arcának. (A törzs- és végtagizmok részvétele ritkább. A hajlító nyakizmok gyengesége gyakoribb, mint a feszítőké, ugyanakkor a csukló- és ujjfeszítők részvétele a folyamatban gyakoribb, mint a hajlítóké.)

A száj-garat izomcsoportok gyengesége miatt megváltozik a beteg által produkált hang, nehézkessé válik a rágás és a nyelés. A hang naziális színezetet kap, amely különösen hosszabb beszédletés során fokozódik. Előfordul, hogy a szájpadizmok gyengesége miatt étkezésnél a folyadék az orron keresztül távozik. A gégeizmok érintettsége rekedtséget okoz. Ez különösen szembeötlő, ha a beteget megkérjük, hogy az elől képzett [ε] hangot hosszan ejtse. Fulladozó, fojtogatásra emlékeztető, torokköszörülő, köhögő hang sem ritka étkezés után, amit a beteg hosszabb idő óta panaszolhat. Hosszan tartó fennállásnál, avagy súlyosabb esetekben jellegzetes arckifejezés figyelhető meg. Fáradt állapotban a száj szöglete lekonyul, depressziós kifejezést kölcsönözve az arcnak. Mosolygásnál a felső ajak közepe felfelé mozdul, a szájzugok horizontális összehúzódása pedig természetellenes vicsorgó szájbígyesztésben nyilvánul meg. Az állkapocs gyengeségét a vizsgáló a kezével észlelheti: a száj ellenállás nélkül nyitható; ez norma-

lis esetben, akaratlagos ellenállásra felszólításkor nem következhet be. Az állkapocs gyengeségének gyakori jeleként a beteg gyakran hüvelykujját az állához, középső ujját ugyanakkor az orra alá vagy az ajkához érinti, mutatóujját pedig az arcához szorítja, mintegy figyelmes kifejezést kölcsönözve tekintetének (Horváth 2001, Horváth–Pataki 1985).

A súlyos izomgyengeség (myasthenia gravis) diagnózisa

A betegség felismerése sok esetben meglehetősen nehézkes; ezért a felmerülő gyanú esetén többféle eljárást, tesztet kell alkalmazni a biztos diagnózis felállítása érdekében. Myasthenia gravisban intravénás edrophonium clorid (Tensilon) hatására az ideg-izom működés jellegzetesen javul, a gyenge szemizmok és garatizmok megerősödnek. A vizsgáló az injekció előtt a beteget maximális izomerő kiváltására szólítja fel az érintett izom területén, majd ugyanezt megismétli a vizsgálat után. Feltűnő javulás figyelhető meg szemhéjcsüngés vagy nazális színezetű beszéd esetén. (Az injekciót részletekben kell beadni, s pozitív esetben az izomerő változása 60 másodpercen belül bekövetkezik.) Saját gyakorlatunkban ezt a Tensilon-tesztet – amely 90%-ban pozitív – minden esetben elektrofiziológiai vizsgálattal (sorozatos ingerlés) kapcsoljuk össze (Horváth–Pataki 1985). Ez a diagnózis felállítását tovább erősíti. Meg kell említeni, hogy a teszt MG-ra nem specifikus. Pozitív eredményt láthatunk például motoros idegi betegségekben vagy a III. szemmozgató agyideg (nervus okulomotorius) betegségében is. (Újszülöttnél és csecsemőnél bőr alá adjuk – kis adagban – az injekciót, ilyenkor a hatás – értelemszerűen – elhúzódik. Előfordulhat, hogy a MG jelenléte ellenére nincs reakció, ilyenkor neostigmint adunk az izomba. Ez utóbbi csecsemőknél is hasznos.)

Az AChR-ellenes antitest koncentrációja a szérumban változó, s a betegség súlyosságát illetően nincs kórjóslati jelentősége. A betegek mintegy 10%-ánál nem mutatható ki (szeronegatív típus, l. Sanders et al. 1997), klinikailag mégis hasonló a betegség kórlefolyása. Másfelől enyhe emelkedése más betegség esetében vagy egészséges rokonoknál is előfordulhat. A csecsemőmirigy daganatos elváltozásánál az AChR-ellenes antitest vérszérum titere gyakran magas (szeroológiailag pozitív). Az elektrofiziológiai vizsgálatok során az ismételt ideg-izom ingerlésnél sorozatban adjuk az ingeráramot az erősen megfeszített célizomra (3 Hz-es frekvenciával), majd digitálisan (számítógéppel) ábrázoljuk és mérjük az izom motoros egységének elektromos válaszát. Az így megjelenő motoros egységpotenciál időtartamából, pontenciálkülönbségéből (amplitúdó) következtetni lehet az izom működésére. Súlyos izomgyengeségben a 4. és 5. potenciál amplitúdójának 10%-nál nagyobb csökkenését észleljük (Horváth–Pataki 1985, Magony et al. 2006). A csökkenés gyakoribb a proximalis izmokban. A teszt pozitivitása 60% feletti. A legérzékenyebb diagnosztikai vizsgálat az ún. egy rost elektromiográfia (Sanders et al. 1996). Noha a teszt súlyos izomgyengeségre nem specifikus, mégis csaknem minden betegnél kimutatható.

A súlyos izomgyengeség (myasthenia gravis) klinikai lefolyása

A betegség lefolyása nagyon változó. Egyes esetekben a gyengeség gyorsan áttér az egyik izomcsoportról a másikra, máskor évekig változatlan a kiterjedése. Megmagyarázhatatlan javulással is lehet találkozni az esetek felében. A betegséget sokféleképpen csoportosították (melyik életkorban kezdődött; a csecsemőmirigy megnagyobbodása, avagy ennek hiánya; AChR-antitest jelenléte vagy hiánya alapján stb.). Saját gyakorlatunkban az amerikai Myasthenia Társaságnak (MG Foundation of America) az Osserman-féle (1958) beosztásra épülő legújabb (2003) klinikai klasszifikációs ajánlását használjuk. Ezt az alábbiakban foglaljuk össze.

I. A szemizmokat érintő súlyos izomgyengeség (okularis MG; előfordulás: 15-20%).

II. a) Enyhe fokban általánossá váló forma lassú romlással. Gyógyszerre jól reagál, nincs krízisveszély (előfordulás: 30%).

II. b) Közepesen súlyos esetek, törzs- és végtagizmok, valamint agytörzsi beidegzésű izomérzékenységgel. Krízisveszély nincs. Gyógyszerre szerényebben reagál (előfordulás: 25%).

III. Heveny, robbanásszerűen súlyosbodó myasthenia gravis, légzőszervi krízissel. Szerény gyógyszerhatás, gyakori a thymoma, magas a halálozási arány (előfordulás: 15%).

IV. Késői típusú myasthenia gravis, mely a III. típushoz hasonló, az I. és II. típus két éven belül progressziójaként alakul ki (előfordulás: 10%).

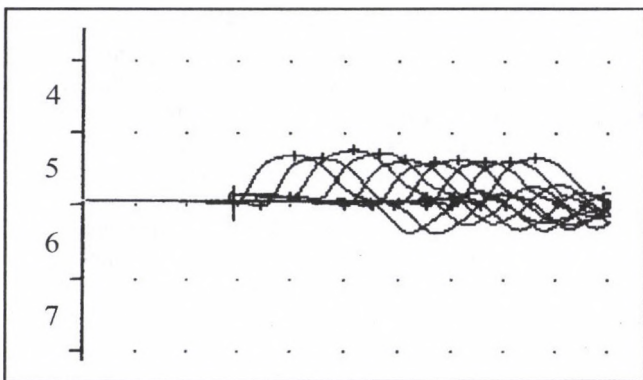
A súlyos izomgyengeségben előforduló krízis (myastheniás krízis) órák alatt kifejlődhet, ezt légzési elégtelenség és a négy végtag bénulása jellemzi. A légutak fertőzése vagy bizonyos gyógyszerek – egyes antibiotikumok, fekélybetegség gyógyszerei, antivirális szerek, nonszteroidok, szívgyógyszerek, vérnyomáscsökkentők, hormonok, koleszterincsökkentők –, amelyek blokkolják az idegizom átvetést, elősegítik a krízis kialakulását. Nemritkán a száj-garat izomzat gyengesége (oropharyngealis MG) miatt kialakult félrenyelés (mint a 2. esetünk-nél) vezet krízishez: a végtagok és a rekeszizom működésének gyengüléséhez.

Az események felgyorsulásának oki tényezője gyakorta nem bizonyítható. Előfordul, hogy egy légzőszervi betegség antibiotikus kezeléssel történő megállítása indítja el a folyamatot. A krízis a MG diagnózisának felállítását után bármikor következhet, de az esetek felében másfél éven belül. A fenyegető állapot első tünete lehet a paradox (hasi) légzés megjelenése, a mellkasfal nagy amplitúdójú mozgása, a légzés ún. vitálkapacitásának csökkenése, szorongás, remegés stb.

A betegség hosszú lefolyású, az első évében gyakoribbak a halálesetek. A második veszélyes periódus a betegség kezdetétől számított negyedik és hetedik év közé esik. A halálesetek főleg légzőszervi komplikációk (fertőzés, félrenyelés és fulladás) miatt következnek be. Míg korábban a halálozási arány a betegség első évében megközelítette a 30%-ot, manapság megfelelő kezelés mellett ez 5% alatt van, és a legtöbb beteg produktív életet élhet.

Esetelemzések

1. K. J. 26 éves nő. A fiatal, tisztjelölteket oktató hivatásos katonanő úgy észlelte, hogy hosszabb előadás után **beszéde elhalkul, orrhangzóssá** válik. Tanítványai ezen jól mulattak, abban a téves hitben, hogy a kezdő tanár zavarban van a nálánál csak néhány évvel fiatalabb ellenkező nemű tanítványai előtt. Sajnos ezt az elképzelést kollégái és orvosok is támogatták, a tünetet „lelki okkal” magyarázták, és nyugtatók szedését ajánlották. A beteg azonban következetes maradt. Határozottan állította panaszainak romlását: az artikulációs és mimikai izmok gyors elfáradását, elmondása szerint ilyenkor a nyelése is nehezebb volt. Érdekességgént említette, hogy előrehajlásnál tünetei fokozódtak. Pihenés után beszéde ismét tiszta, érthető volt. Pajzsmirigyműködése fokozott volt, az utóbbi két évben 13 kg-t fogyott. Ismételt vizsgálatok a súlyosbodó beszédpanaszt munkahelyi körülményeiből eredő neurotikus tünetképződéssel, szorongással, pánikrohammal, másrészt pajzsmirigy-túlműködéssel magyarázták. Neurológiai vizsgálattal organikus eltérésre utaló tünet nem volt észlelhető. A fárasztási teszt értékelése sem tűnt meggyőzőnek. Ismételt idegingerléssel végzett elektrofiziológiai terhelési tesztje enyhén pozitív lett (2. ábra). A vénás Tensilon (edrophonium chlorid) adására kialakult kolinergia csak mérsékelt reakciót (fokozott nyálképződés, könnyezés) mutatott, továbbá a gyógyszerhatás alatt mért terhelési értékek sem voltak bizonyító erejűek.

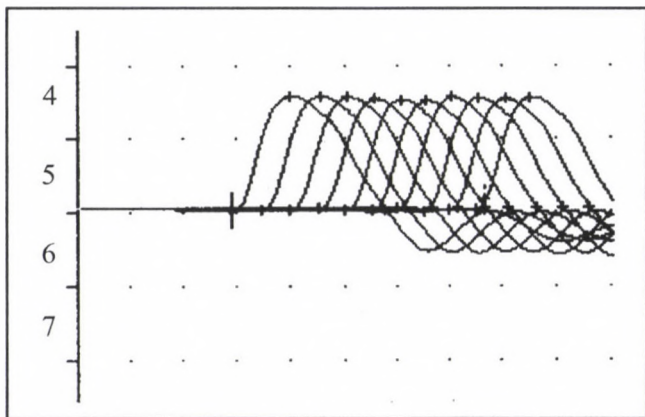


2. ábra

Vázizomzati EMG-vizsgálat során a sorozatos ingerlés eredménye
myastheniás fiatal nőbetegnél:
az 5. és 6. potenciál amplitúdója csökkenést mutat

A neurológiai osztályon tartózkodása alatt kiderült, hogy 8 hetes terhes. Kombinált terheléssel és edrophoniumtesztrel kiegészítő vizsgálatokat végeztünk, amelyek során az ismételt (repetitív) stimuláció egyértelmű pozitivitást mutatott (3. ábra). Az ezt követő antitesttiter a myasthenia gravis diagnózisát erősítette

meg. Az a tény, hogy a mellkasi CT-n a csecsemőmirigy megnagyobbodása volt látható, igazolta az MG feltételezését.



3. ábra

Vázizomzati Tensilon-tesztel kombinált EMG-vizsgálat során az ismételt ideg-izom ingerlésre egyenletes amplitúdójú motoros egységpotenciálok láthatók ugyanannál a fiatal nőbetegnél

A továbbiakban alaposan mérlegelni kellett a várható szülés és az esetleges állapotrosszabbodás kockázati tényezőit. Mivel a beteg állapota fokozatosan romlott, és egyre nagyobb dózísú Mestinont igényelt, felelőtődött a műtét lehetősége. A szövettani vizsgálat a csecsemőmirigy túlbujánzását mutatta, rosszindulatú sejtosztódás jele nélkül. Ezután a beteget megműtötték, aminek során a szegycsont átvágását követően teljes csecsemőmirigy-eltávolítást végeztek. A műtétet követő harmadik naptól szempanasz jelentkezett; a tárgyakat elmosódva látta, beszéde elkennté vált, karjait nem tudta megemelni. Úgy tűnt, hogy betegsége újabb izomcsoportokra (szem- és vázizomzat) terjedt ki. Ez az átmeneti állapot műtét után gyakori, s a légzőizmok elgyengülésének veszélye miatt lélegeztetőgép közelségét igényli. K. J. panaszainak súlyossága a kezelésként adott Mestinon adagjától függően hullámzott, majd pár hónapon belül fokozatosan panasz- és tünetmentessé vált, s fél év múlva a gyógyszerét is elhagyhatta. A műtétet követően két hónappal császármetszéssel egészséges újszülöttnak adott életet, akit a boldog szülők büszkén mutattak be az orvosoknak. Volt betegünket három év óta rendszeresen ellenőrizzük; panasz- és tünetmentes, anya és gyermeke jól vannak.

2. J. S. 60 éves, enyhe magas vérnyomásban és cukorbetegségben szenvedő férfi panaszai 2005 nyarán csak szűk környezete által észlelt **orrhangzós beszéddel** (rhinophoniával) kezdődtek. Ez később kisfokú nyelészavarral és légzészavarral társult. A területi szakrendelőben kivizsgálták, és kétoldali kis epi-

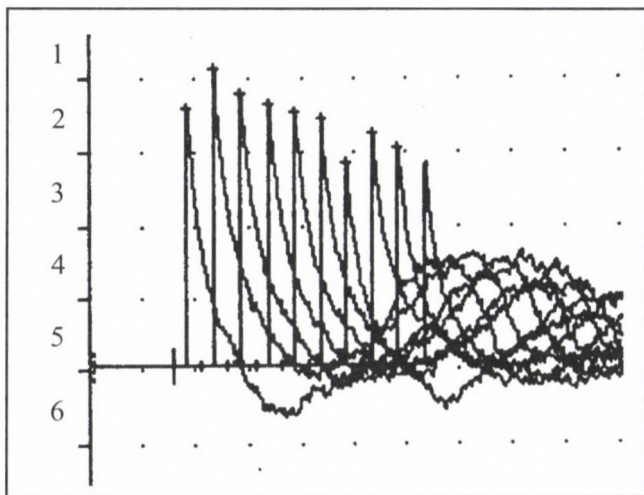
glottiszcsisztát találtak nála, amelyet altatásban megoperáltak. A műtét után súlyos szövödményként aspirációs tüdőgyulladás alakult ki; a beteg nem ébredt fel, és a heveny légzési elégtelenség tünetei miatt komplex intenzív ellátást igényelt gépi lélegeztetéssel. Állapota nem javult, ezért átmenetileg gégemetszésre (tracheosztómára) szorult, majd pár hét múlva erőteljes antibiotikus kezelés hatására tünetei fokozatosan javultak, enyhe tünetekkel hazabocsátották.

10 nap múlva, kórházi kivizsgálása során hirtelen fulladásos rosszullét jelentkezése miatt ismételten félrenyelés lehetősége merült fel, emiatt altatásban mikrolaringoszkópiás vizsgálat történt, amely után – az ébresztést követően – ismét akut légzési elégtelenség alakult ki nála. Ezért újra 24 órás gépi lélegeztetésre szorult, ami után légzése kompenzálódott, járóbetegként kezelték tovább. Röviddel ezt követően hirtelen légzésleállás és keringési elégtelenség tünetei miatt intenzív osztályra került, harmadszor is gépi lélegeztetést kellett alkalmazni a betegnél. Az akkor végzett akut koponya-CT nem mutatott kóros eltérést. Néhány nappal később kétoldali szemhéjcsüngés (ptosis), diszfónia és súlyos nyelészavar jelentkezett nála. A bulbaris tünetek felvetették az agyi történés, az agytörzsi szindróma, illetve az amiotrophiás lateralsclerosis (ALS = felső motoros idegbetegség) lehetőségét is. Emiatt további vizsgálatok céljából egy másik kórház intenzív osztályára került. A végtagok izmaira, illetve a törzsre kiterjedő (ún. szkeletális, vázizomzati) tünetei enyhe fokúak voltak. Újabb néhány nap múlva rohamkocsival, asszisztált gépi lélegeztetéssel szállították a Rókus Kórház Neurológiai Osztályának EMG-laboratóriumába (Tensilonnal kombinált velopharyngealis és vázizomzati EMG elvégzésére). A sorozatos ingerlés hatására az amplitúdó fokozatosan csökkent.

A klinikai vizsgálat során ALS-t nem tudtunk valószínűsíteni; ugyanis ennél a motoneuron betegségénél legtöbbször csak a folyamat lefolyásának késői fázisában alakul ki a légzőizmok gyengesége, amikor már az izomsorvadás (izomatrophia) kifejezett jelei is megfigyelhetők. A vizsgálat elvégzése után 1 perccel a beteg beszéde érthetővé vált, ptosisa javult, a felkínált folyadékot meg tudta inni. A vázizomzat ismételt ingerlésére adott amplitúdócsökkenés válaszreakciója 10%-os volt, ami diagnosztikus értéket jelentett (4. ábra). A válaszreakció edrofónium adására normalizálódott.

A beteg mellkasi CT-je kóros mediasztinális eltérést nem mutatott. Az ismételt hörgőtükrözés és sejtvizsgálat (bronchoszkópia, citológia), illetve szövettani vizsgálat a bal hörgőknél 2 mm-es makroszkópos lumenbedomborodást mutatott, de szövetátalakuláson (hám-metaplázián) kívül rosszindulatú szövettani elváltozás jelenlétét kizárta. Nehézségek jelentkeztek a diagnózis felismerését követő kezelés során is. Először Mestinont kapott, de kolinerg mellékhatás tüneteinek (fokozott hörgőváladékozás, extrém nyáleválasztás, hasmenés) jelentkezése miatt az adagot csökkentettük, majd végleg el kellett hagyni. Ezután ötször (kezdetben másnaponta, majd harmadnaponta) plazmaferézis kezelést és 125 mg véna-s kortikoszteroidot kapott. Emellett klinikailag az agytörzsi tünetek hullámzó intenzitása volt észlelhető. A gépi lélegeztetésről néhány héttel később többször

próbálták levenni agresszív légzőterápia, masszázs segítségével, de ez meglehetősen nehezen sikerült. A kialakult pánikszindróma miatt a gégekanült sokáig nem lehetett megszüntetni. Táplálását az intenzív osztályon átmenetileg PEG (hasfalon átvezetett tápszonda) beültetésével oldották meg. Jelenleg a beteg tünetmentes, jól van, járóbetegként ellenőrzik.



4. ábra

Vázizomzati EMG. Agytörzsi eredetű myasthenia gravisos betegnél elvégzett ismételt ingerléskor 10%-os amplitúdójú csökkenés látható

Következtetések

Beszédünk nem állandó; átmeneti és fokozatosan létrejövő változások jellemzik. A hirtelen fellépő, hallásra is feltűnő eltérések azonban normális esetben nem jellemzik az emberi beszédprodukción. A nazálissá, rekedtté váló, kellő ok nélkül gyorsan fáradó hang, a nehézkes beszéd mind-mind súlyos betegség tünetei is lehetnek, ezért kellő odafigyelést igényelnek. Noha nem áll rendelkezésre minden emberről meghatározott életkoronként egy olyan beszédminta, amelyhez – a panaszok fellépése esetén – hasonlítani lehetne a kórosnak tűnő beszédet, a szubjektív benyomások beszédakusztikai vizsgálatokkal alátámaszthatók. Ezek a korszerű elemzések pedig objektív tényeket adhatnak a neurológusnak az időben felállítandó helyes diagnózishoz. A myasthenia gravis mint betegség lehetőségét igen sok esetben a logopédus, illetve a klinikai fonetikus diagnosztizálja más országokban is (Duffy 2006). Meggyőződésünk, hogy a jövőben az interdiszciplináris területek művelői fokozott mértékben kell, hogy együtt dolgozzanak a közös cél érdekében.

Irodalom

- Bradley, Walter G. – Daroff, Robert B. – Fenichel, Gerald M. – Jankovic, Joseph (eds.) 2000. *Neurology in clinical practice*. Elsevier, Amsterdam.
- Duffy, Joseph R. 2006. Distinguishing among motor speech disorders is important: The role of speech pathology in neurologic diagnosis. 5th International Conference on Speech Motor Control, Nijmegen. Abstracts. *Stem-, Spraak- en Taalpathologie* 14. 117–118.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Gósy Mária 2005. *Pszicholingvisztika*. Osiris Kiadó, Budapest.
- Horváth Szabolcs (szerk.) 2001. *A kommunikáció fejlődése és idegi eredetű zavarai*. Animal-Med Kft., Budapest.
- Horváth Szabolcs – Pataki László 1985. *A nyílt orrhangzóság és a diszfónia neurofoniátriai vonatkozásai*. Kandidátusi értekezés. Budapest.
- Magony Edit – Horváth Szabolcs – H. Fehér Éva 2006. A diagnózis nem könnyű, de gondolni kell rá, mert azonnal életet ment. Krízis MG-ban. Előadás. Pest Megyei Orvosnapok, Gödöllő.
- Sanders, Donald B. – Schilberg, Erik V. 1996. AAEM minimonograph 25: Single-fiber electromyography. *Muscle & Nerve* 19. 1069–1083.
- Sanders, Donald B. – Andrews, P. I. – Howard, James F. – Massey, Janice M. 1997. Seronegative myasthenia gravis. *Neurologia* 48. 540–545.
- Szobor Albert 1977. A myasthenia gravisről. In Sárkány Jenő (szerk.): *Gyermek-neurológia*. Orvostovábbképző Intézet, Budapest.
- Szobor Albert 1983. *Myasthenia gravis*. Doktori disszertáció. Budapest.
- Thomas, C. E. – Mayer, S. A. – Gangor, Y. 1997. Myasthenic crisis clinical features, mortality, complications, and risk factors for prolonged intubation. *Neurology* 48. 1253–1260.

ÚJ RENDSZERŰ, KORPUSZALAPÚ GÉPI SZÖVEGFELOLVASÓ FEJLESZTÉSE ÉS KÍSÉRLETI EREDMÉNYEI

Németh Géza – Olaszy Gábor – Fék Márk

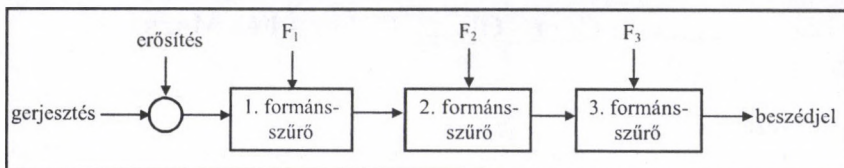
Bevezetés

A beszéd gépi előállítás ma már nem különlegesség, hiszen a mindennapi életünket körülvevő információközvetítő eszközökben, rendszerekben gyakran alkalmazzák. A géppel generált beszéd minősége változhat attól függően, hogy milyen információt kell beszéddé alakítani. Minél inkább kötött a tematika, ebből következően a szintetizálendő mondatok szerkezete, annál jobb minőséget lehet elérni. Emellett a technikai lehetőségek is behatárolhatják, hogy milyen módszert alkalmaznak a beszéd előállítására (például mobiltelefonban szól, vagy egy központi szerver generálja a beszédjelet). A kutatók célja, hogy minél jobb minőségű gépi beszédet állítsanak elő. Az eddig kifejlesztett módszerek egyre jobban érthető beszédet szolgáltatnak, a minőség is közeledik az emberi hangzáshoz (nem robotos), azonban például a beszélő egyéni hangszínezetét, kiejtési stílusát még ritkán lehet felismerni a gépi szövegfelolvasóknál. Hangszínezeten értjük a személy egyéni hangját (voice timbre) normál beszédben, amely alapján például egy ismerőse felismeri a beszélőt. A korpuszalapú beszéd szintézis módszere e legutóbbi hiányosságot is teljesíteni tudja, ezért várhatóan egyre jobban terjedni fog. A tanulmányban bemutatjuk az első, Magyarországon fejlesztett ilyen rendszer kísérleti eredményeit, ugyanakkor rávilágítunk arra is, hogy milyen korlátokkal kell számolni.

A beszéd szintézis korábbi módszerei

Minden szövegfelolvasó rendszer elméletileg két alapvető részből épül fel. Az első rész a bemeneti szöveget értelmezi, és szimbolikus információvá alakítja, a második a szimbolikus információt alakítja át beszéd-hullámformává (általában valamilyen hangfájlt állít elő). A közbenső szimbolikus információ általában a szöveg tartalmát megadó fonémasorozatból és a beszéd prozódiai jellemzőit (hanglejtés, hangsúlyok, ritmika) leíró információkból áll össze. A megoldási módszerek között abban vannak eltérések, hogy az előbbi két részt megvalósító belső megoldások milyen elvek szerint történnek. Az alábbiakban röviden sorra vesszük az eddigi beszéd szintézis-technológiák főbb jellemzőit. Megjegyezzük, hogy a technikai lehetőségek fejlődése döntően befolyásolta a módszerek kialakulását is.

A **formánsszintézis** volt az első olyan technológia, amelynek segítségével egy szöveget automatikusan beszéddé lehetett alakítani. Az eljárást napjainkban is alkalmazzák. A formánsszintézissel az ember beszédképzési folyamatát utánozzák (gerjesztő jel + toldalékcső) elektronikus formában (gerjesztett, dinamikusan hangolt szűrőrendszer). A formánsszintetizátor egy lehetséges megvalósítását az 1. ábra mutatja.



1. ábra

Soros elrendezésű formánsszintetizátor blokkvázlata

A gerjesztés zöngés hangoknál a hangszalagok által keltett kváziperiodikus jelnek feleltethető meg, illetve zöngétlen hangok esetén zajszerű. Egy-egy szűrő a megadott formánsfrekvencia környezetében erősíti a gerjesztés felharmonikusait, ezzel modellezve a garat, a gége és a szájüreg által alkotott rezonátorrendszer erősítéseit. A formánsfrekvenciák definíciószerűen a zöngés hangokat jellemzik, de zöngétlen gerjesztésnél is felhasználhatók a zörejes gerjesztésű hangok jellemző zörejtőcainak az előállítására (ez műszaki egyszerűsítés). Az első három formánsfrekvencia jól leír egy-egy beszédhangot. A formánsok adatait (a szűrők rezonanciafrekvenciáit) legalább 10 ms-onként meg kell adni ahhoz, hogy érthető beszéd álljon elő. Az adatok megadását szabályrendszer vezérli, amelyet általában fonetikusok határoznak meg. A formánsszintézis során a beszéd prozódiaját is szabályrendszerrel írják le, majd a szabályok alapján változtatják az alapfrekvenciát (a beszéddallam, illetve a hangsúlyozás kialakítására), az intenzitást (a hangsúlyozás, illetve a hangerő beállítására) és a hangok időtartamát (a ritmus megvalósítására). A formánsszintetizátort tehát úgy jellemezhetjük, hogy a beszéd szegmentális és szupraszegmentális részét egyaránt szabályok alapján állítja elő. A gyakorlat azt mutatja, hogy ez a módszer ugyan érthető beszédet szolgáltat, de nem adja vissza az emberi hangszínezet finomságait (robotos a hangzás). A módszer előnye a kis tárcapacitás- és az alacsony számításigény. A BME Távközlési és Médiainformatikai Tanszékén (TMIT) kifejlesztett Multivox magyar nyelvű formánsszintetizátor (Olasz et al. 1992) ingyenesen hozzáférhető.

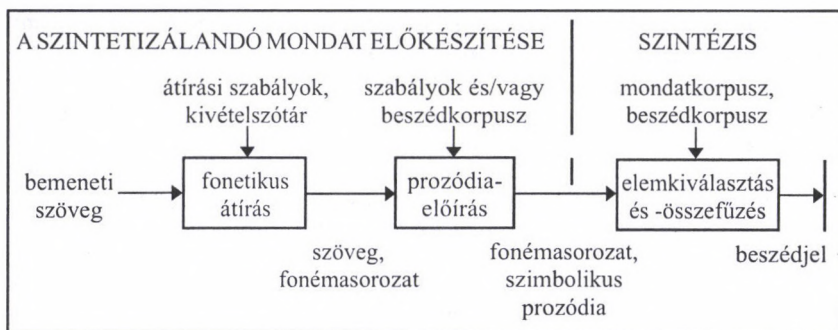
Az **elem-összefűzéses** módszert a beszéd minőségének javítására dolgozták ki. A ma működő gyakorlati alkalmazások többsége ezen az eljáráson alapul. A döntő különbség a formánsszintetizátorhoz képest az, hogy a beszédhangok hullámformájának előállítási módszere változott. A prozódiai részt itt is szabályokkal modellezik. Az eljárás lényege, hogy a formánsfrekvenciák időről időre történő megadása helyett visszatértek az emberi beszédhez. A számadatok helyett az em-

ber által felolvasott beszédből kivágott, kész hullámformákat tárolják el, majd ezeket fűzik össze. A gondolat lényege, hogy ezekben a hullámformarészletekben automatikusan benne vannak a formánsok frekvenciaértékei, azok változása a ko-artikulációban, továbbá még a gerjesztést sem kell parametrikusan megadni. Elvárható volt tehát, hogy ha ilyen hullámformaelemeket összefűzünk, akkor olyan hangszínezetet kapunk, amely közelebb áll az emberi beszédhez, mint a formáns-szintetizátor hangja. Ezért nevezték el a módszert elem-összefűzéses technikának. A technológia egyik alapvető kérdése, hogy melyek legyenek azok a hullámformaelemek, amelyek összefűzésével létrehozzuk a gépi beszédet. Úgy kell tervezni, hogy az adott nyelv tetszőleges hangsorozatát elő tudjuk állítani. A gyakorlatban bevált kompromisszumos megoldás a két egymással összekapcsolódó félhang együtteseként előálló diád hullámformájának a tárolása. A magyarban 1600 diáddal lefedhető a teljes hang- és hangkapcsolódási állomány (Olaszy 1999). A szupraszegmentális szerkezet paramétereit ugyanakkor itt is szabályrendszerrel határozzák meg, és jelfeldolgozással végzik az alapfrekvencia-, az intenzitás- és a hangidőtartam-változtatásokat. Az eljárás előnye a szebb, emberibb hang, továbbá az is, hogy a rendszer működésének minden részlete jól kézben tartható (ez fontos a hibakeresésnél). Hátránya, hogy a hangokat a közepükön fizikailag elvágjuk, ami torzítást okoz, főként a zöngés hangok esetén. A torzítás csökkentésére alakították ki a triádós hullámformát, mint másik alapelemet, amelyben egy félhang + egy teljes hang + egy félhang jelent egy-egy tárolt elemet. Így a középső hangban nincs vágás, tehát a torzítás csökken. A legtöbb diádós, diádós és triádós hullámforma-elemtára jellemző, hogy minden elem csak egyszer fordul elő (zárt szerkezetű elemtár), így az összefűzésnél nem lehet az elemtárban lévő elemekből válogatni (ha a hangsor több pontján van szükség ugyanarra az elemre, akkor az minden esetben ugyanaz lesz, belső tartalma nem változik). A zárt elemtárral a finom beszédrészleteket nem lehet megvalósítani, ugyan az egyénenkénti hangkarakter már részben hallható (más bemondótól készített elemtár más hangot eredményez), azonban a beszélő egyéni hangszínét továbbra sem tudják visszaadni. Ennek oka egyrészt a zárt szerkezetű elemtár, másrészt az, hogy az egymás mellé illesztett beszédhullámformák (diád, triád) túl rövidek ahhoz, hogy az egyéni jellegzetességek karakterisztikusan érvényesüljenek. További ok, hogy a prozódia továbbra is az egyéni jellemzőket figyelmen kívül hagyó szabályrendszer alapján áll elő. A BME TMIT-en kifejlesztett Profivox magyar nyelvű beszédsszintetizátornak (Olaszy et al. 2000) létezik tiszta diádós elemtárral működő változata, illetve 1444 diádból és 6000 triádból álló nagyobb elemtára. A kétféle rendszer hangzásában (ugyanazon bemondótól származó elemtárral) a triádok alkalmazása minőségi javulást eredményez (Nagy et al. 2005).

A korpuszalapú beszédsszintézis

A korpuszalapú beszédskeltés adja napjainkban a legjobb minőségű szintetizált beszédet (Schweitzer et al. 2003). Ennél a módszernél a beszélő hangszínezete, kiejtési stílusa egyértelműen felismerhető. A technológia nevéből adódik, hogy

egy adott, a beszédhangok változatossága szempontjából elméletileg nyitottnak tekinthető **beszédkorpusz** helyettesíti a korábban alkalmazott zárt szerkezetű, diádok, triádok elemtárat. A nagy korpusz sok órányi beszédet tartalmaz, annak hullámformaszintű tárolásával és címkézésével. A beszédkorpuszal párhuzamosan tároljuk annak szöveges formáját is, ezt nevezzük **mondatkorpusznak**. Ebben végezzük a keresést, majd, ha megtaláltuk a megfelelő szövegrészt, azt kiemeljük a beszédkorpuszból. Az így kiemelt hullámformákat összefűzzük, így áll elő a szintetizált mondat hullámformája. A korábbi rendszerek módszertanához képest ez nagy változás. A formánsszintetizátoroknál döntően fonetikai tudásra volt szükség a vezérlő paraméterek meghatározására mind szegmentális, mind szupraszegmentális szinten. Az elem-összefüztetés technológiánál a hangsorszerű összeállításnál már visszakanyarodtak a kutatók az emberi beszédhez (diádok, triádok tárolása), és csak a prozódia megvalósításához használtak konkrét fonetikai ismereteket. A korpuszalapú rendszereknél a hangsor-összeállítás-hoz igen nagy méretű, nyitott beszéd- és mondatkorpuszt használnak, és a szegmentális és szupraszegmentális szerkezetre vonatkoztatott összegzett fonetikai tudást áttelesen építik be az elemkereső és összefűző költségfüggvénybe. A módszer mögötti mérnöki gondolat a következő. Egy bemondó által felolvasott mondat tartalmazza a hangsor hullámformáját, és ebben benne foglaltatik a prozódia is. Ha sok mondatot tárolunk az egyébként nyitott beszédkorpuszban, akkor nagy a valószínűsége annak, hogy egy szintetizálendő mondat előállítására a korpuszban találunk olyan szavakat, szókapcsolatokat, mondatrészeket, amelyek a legjobban illeszkednek (hangsorilag és prozódiaiilag) a készítendő beszéd adott pontjához. A mondatot így viszonylag hosszú egységekből, szavak, szókapcsolatok hullámformáinak sorozatából össze lehet állítani (2. ábra).



2. ábra

Korpuszalapú beszédszintetizátor bloksémája

Ha a szavak kiválasztásánál a prozódiai tartalmat is figyeljük, és csak a megfelelő prozódiajú szót engedjük kiválasztani a beszédkorpuszból, akkor a szintetizált mondatban a prozódia is meg fogja közelíteni az optimálist. A két feltétel

teljesítéséhez – elsősorban a szókapcsolások folytonosságának biztosításához – a szóvégi és szókezdő hangok szerinti fonetikai szabályrendszert kell alkalmazni. Ez megmondja a válogatásnál, hogy két összekapcsolandó szó találkozási pontján a hangok spektrálisan (pl. formánsmenetben) illeszkednek-e. Modellezni kell ezenfelül a mondat prozódiaját is. Ha ugyanazt a modellt használjuk a beszéd-korpuszban, a mondatkorpuszban és a szintetizálendő mondatban is, akkor a prozódiai tartalom szempontjából is tudunk keresni a beszédkorpuszban, el tudjuk dönteni, hogy a hangszinten összekapcsolhatónak ítélt szó, szókapcsolat prozódiailag is megfelelő-e. Ha igen, akkor alkalmazzuk, ha nem, akkor tovább keresünk. (A beszédkorpusz mondatainak sokasága elméletileg nincs korlátozva, azonban a feldolgozási sebesség és az optimális keresési módszer behatárolja ezt a szabadságot.) A hipotézis tehát az, hogy a beszédkorpuszból való gondos összeválogatás és egymáshoz kapcsolás esetén az előállított szintetizált beszéd minősége igen közel lesz a korpusz eredeti beszédének a minőségéhez. A fentiek megvalósítását lehetővé tette a számítástechnikai eszközök rohamos fejlődése memóriakapacitásban és feldolgozási sebességben. Jó példa erre egy Japánban készített rendszer, amelynek beszédkorpuszát 80 órás hangfelvételtől alakították ki (Kawai et al. 2004). Az ilyen rendszerek legfőbb előnye, hogy a hangminőség nagyon jó lehet (a szintetizált mondatot össze lehet téveszteni az eredeti bemon-dó hangjával). Hátránya, hogy nehéz az összefüzendő elemeket kiválasztó függ-vényt optimálisra megtervezni, hogy hatalmas adatbázisokkal kell dolgozni, hogy nehéz a hibakeresés, valamint a javítás. Ezek miatt első lépésben csak kö-tött, jól körülhatárolt témában van esély az optimális eredmény elérésére. A fen-tieket figyelembe véve kísérleti rendszerünkkel csak időjárás-jelentések jó mi-nőségű előállítását tűztük ki célul. A módszerrel előállított beszédjel természe-tessége, így hangminősége ugrásszerű javulást mutat a korábbi módszerekhez képest. Ezt az teszi lehetővé, hogy a diád és triád elemeknél hosszabb, folytonos beszéddarabok kerülnek egymás mellé, megőrizve azok eredeti hangszínezetét, ritmikai tartalmát, hanglejtését. Továbbá a nagy beszédkorpusz lehetővé teszi, hogy spektrálisan is, és prozódiailag is egymáshoz jobban illeszkedő beszéd-sza-kaszokat válogassunk és fűzzünk össze, ami szintén a természetes hangzást biz-tosítja. Ez a módszer közelíti meg legjobban a biológiai beszédkeltést, azt, hogy **az emberi beszéd egyedi és egyszeri produktum**. Az egyedi jelző az egyén sa-ját hangjára vonatkozik, az egyszeri pedig azt fejezi ki, hogy a beszédjelet az adott időpillanatra jellemző biológiai rendszer hozza létre (mindkettő pillanatnyi állapotától függően). A beszédprodukciónak akusztikuma tehát a megvalósulási idő-tengelyhez kapcsolódik (gondoljunk arra, hogy sokórás hangfelvétel képezi a beszédkorpuszt, ennek az időtengelyén kell keresni). Ha ebből az időtengelyből viszonylag hosszú szakaszokat ragadunk ki (szó, szókapcsolat), akkor a hangzás hordozni fogja a beszélő hangszínezetét. A hosszabb beszédrészlet jobban meg-közelíti az egyszeri és egyedi ejtésre jellemző hangzást, mint a rövid. A korpusz-alapú szintézisnél az elérhető minőségjavulást a válogatási tér, azaz a korpusz nagysága befolyásolja. Az ideális az lenne, ha a korpusz tartalmazná a szinteti-

zálandó mondatok mindegyikét. Ezt nem lehet megvalósítani, ezért a fejlesztők arra törekszenek, hogy a műszaki lehetőségekhez optimalizálják a beszédkorpusz nagyságát és tartalmát.

A szintézis alapegysége és a prozódia megvalósításának egy lehetséges modellje

A korábbiakból már láthattuk, hogy a beszéd szintézisnél a beszédjelet minden esetben valamilyen jól definiált részegységekből rakjuk össze. A beszédkorpusz készítésénél is az első alapkérdés, hogy mi legyen ez a jól definiált alapelem. Mivel nyitott szerkezetű beszédkorpuszról van szó, elvileg bármilyen hosszúságú hangelem lehet (hang, diád, triád, szótag, szó, mondat stb.) a jelölt, azt a tervező határozza meg. A választástól függ a további korpusztervezés menete. Szónál hosszabb alapelemet nem célszerű használni, mert nehéz a komplett lefedést biztosítani (hogy minden lehetséges hangsort elő lehessen állítani). A szónál rövidebb elemeket pedig már kipróbáltuk a korábban fejlesztett rendszerekben, azok előnyeit, hátrányait ismerjük. Marad a szó. A szó választása természetes hangminőséget jelenthet, hiszen olyan nyelvi egységet határoz meg (legyen rövid vagy hosszú), amelynek észlelésére perцепиós rendszerünk is fel van készítve. A szó kiejtésében automatikusan benne van a szegmentális és a pillanatnyi szupraszegmentális szerkezet is (még ha a szavak összefolynak is a folyamatos beszédben). A jó hangminőség még fokozható azzal is, ha a szintetizálendő mondatot sikerül szókapcsolatokból felépíteni (találunk olyan, több szóból álló mondatrészeket a beszédkorpuszban, amelyeket közvetlenül felhasználhatunk a szintetizálendő mondat felépítéséhez). Ilyenkor még hosszabb hangsorrésze vonatkozik az előbbi állítás. Mindezekből adódóan a most bemutatott rendszerben a keresés alapelemének a szó nagyságú egységet választottuk. Emellett azonban gondoskodni kell egy alacsonyabb szintű építőelemről is, arra az esetre, ha nem találunk megfelelő szót a keresés során. Ez a tartalék elem pedig a hang (hangonként is össze lehet állítani a szükséges szót). Tehát a hangok is építőelemeknek számítanak a rendszerben, csak igen ritkán használjuk őket. Mindezekből következik, hogy a beszédkorpuszban minden mondat hullámformájába be kell jelölni minden hang határát és minden szóhatárt (lásd később). Vegyünk egy példát. Egy 10 szavas mondathoz például a legrosszabb esetben 10 szót kell keresni a beszédkorpuszban. Emlékeztetünk arra, hogy a formánsszintetizátoroknál az alapegység a 10 ms-onként megadott adatsorozat volt, a hullámforma-szintézisnél pedig a diád, illetve a triád. Mindkét esetben sokkal több elemből kellett az előbbi mondatot összeállítani, mint a korpusz szintézisnél. A beszédkorpusz készítésének második alapkérdése a prozódiai modell. A kiindulási gondolat fontos alapeleme, hogy a korpuszalapú szintézisnél a prozódia előállítására lehetőleg ne alkalmazzunk mesterséges dallam-, intenzitás- és időtartam-módosítást (mivel az torzíja a hangot), hanem keressük a beszédkorpuszban azt az optimális szót, amelyik prozódiai szempontból is illik a szintetizálendő mondat adott részéhez. A felolvasandó szöveganyag mondatait (mondatkorpusz) tehát úgy kell összeál-

lítani, hogy ugyanabból a szóból prozódiaileg többfajta is legyen a mondatokban. A kérdés az, hogy az előbbi általános megállapítást hogyan lehet modellezni. Hányfajta szóvariánst kell a szövegbe ágyazva tárolnunk egy-egy szóból? A prozódia modellezésében alapegységnek a mondatot tekintjük. A modell szorosán összefügg a szintetizálendő szöveg szerkezetével, jelen esetben csak kijelentő mondatokat modellezünk. A kijelentő mondat prozódiai szerkezete jól körülhatárolható, ismert egységekből áll. Ezeket az egységeket a mondaton belüli hely szerinti pozicionálással (hol van a szó a mondatban), valamint a központozással (vesszők, gondolatjelek stb.) kapcsolatba lehet hozni. Ez a modell lényege. Ugyanazt a modellt alkalmazzuk a mondatkorpuszban, a beszédkorpuszban és a szintetizálendő mondatban is. A mondat- és beszédkorpuszban felcímkézzük a mondatokat a modell szerint, a szintetizálendő mondatra pedig alkalmazzuk a modellt. Így prozódiai vonatkozásban is ki lehet alakítani keresési kritériumokat.

A mondatkorpusz kialakítása időjárás-jelentések automatikus felolvasásához

A mondatkorpusz kialakításához meg kell határozni, hogy az milyen mondatokat tartalmazzon. Belátható, hogy ehhez olyan nyers szöveganyagot kell összeállítani, amelyiknek a szóállománya lefedi a majdan szintetizálendő időjárás-jelentés mondatok szóállományát. Első lépésben ennek a mondatállománynak a gyűjtését végeztük el. Egy éven keresztül gyűjtöttünk (saját, automatikus szoftverrel) magyar időjárásjelentés-szövegeket 20 különböző weboldalról. Az eredmény 56 000 mondat, bennük 493 000 szó és 43 000 szám. A teljes szöveg 5 200 különböző szóalakot tartalmazott (a szóalak akkor különbözik, ha a szó betűkarakteres formájában legalább egy karakter különbözik). A statisztikai analízis azt mutatta, hogy a leggyakoribb 500 szóalak lefedte a mondatállomány 92%-át, 2 300 szóalak pedig a 99%-át (prozódiai szempontok nélkül). Ez a kis szám abból ered, hogy a témakört limitáltuk az időjárásos mondatokra. (Megjegyezzük, hogy korlátozás nélküli szöveg hasonló fedéséhez a tárolt mondatállománynak mintegy 70 000 szóalakot kellene tartalmaznia.) A második lépésben alakítottuk ki az 56 000-es mondatállományból a későbbiekben felolvasásra kerülő mondatkorpuszt (5 260 mondat), amely tartalmazta az 5 200 szóalakot és azok prozódiai variánsait (összesen 82 000 szó). Ez a mondatkorpusz képezi a beszéd szintetizátor elsődleges (szóalapú) keresési terét.

A beszédkorpusz elkészítése

A beszédkorpusz a mondatkorpusz felolvasásából jött létre. A hivatásos női bemondó 4 héten át, heti 2-3 alkalommal, naponta 4-5 órát olvasva mondta fel a mondatkorpusz 5 260 mondatát. Az eredmény 11 órás folyamatos beszédanyag – ez képezi a beszédkorpuszt. A folyamatos beszédanyagot mondatokra daraboltuk, minden mondat kapott egy azonosítót. Ezután a mondatkorpusz szöveges formáját (minden mondatát) manuálisan össze kellett vetni a beszédkorpusz tartalmával, kijavítottuk az esetleges felolvasási hibákat (a bemondó néha tudat

alatt átformálta az írott szöveget, ilyenkor a szöveget a felolvasott formához igazítottuk, továbbá a felolvasott számokat és rövidítéseket is szövegesen ki kellett fejteni (például $4-6^{\circ}\text{C} = \text{négy, hat Celsius-fok}$). A cél az volt, hogy teljes szinkronba hozzuk a hangot és annak szöveges formáját. A következő lépésben minden mondat hullámformáját elláttuk szóhatárokkal, hanghatárokkal és a hangok jeleivel (fonetikusan átírtuk a szövegeket). Ezt a BME TMIT automatikus beszédfelismerő szoftverének (Mihajlik et al. 2002) támogatásával végeztük. Az automatikusan átírt és címkézett anyagot – a szó- és hanghatárokat – félautomata módszerrel ellenőriztük. Itt felhasználtuk például a magyar beszédre kidolgozott időtartammodellt (Olasz 2006), minden hangra jósoltunk egy időtartamot, és összehasonlítottuk a bejelölt értékkel. Nagy eltérés esetén manuálisan megkerestük a hiba helyét, és korrigáltuk a rossz jelzést. A szóhatár jelölése sok esetben nem végezhető el egyértelműen. Ezért külön jelet használtunk a szó kezdetének jelzésére (<), és külön jelet a befejezésére (>). Hangösszeolvadás esetén a szóhatár nem jelölhető ki egyértelműen a fonetikus változatban (például: a *Balatonnál legalább* átírása: <balatoná<I>egaláb>). Ilyenkor az „előző szó vége” jelzést megelőzi a „következő szó eleje” jelzés, és ezt az algoritmus értelmezi. Mindezekből látható, hogy a korpuszalapú szintézis beszédkorpuszának végleges formára hozása többszintű, kitartó munkát igényel.

Elemkiválasztás

Az elemkiválasztás algoritmus a korpuszalapú szintézis legproblematisabb eleme. A bemeneten rendelkezésre áll a szintetizálendő mondat szövege a magyar helyesírás szabályainak megfelelően leírva. Ezzel szemben áll a mondatkorpusz mint keresési tér a szöveges és fonemikus formákkal. Az elemkiválasztó veszi a szintetizálendő mondat első szavát, és elkezd keresni a mondatkorpusz szövegében. A legtöbb esetben sok jelöltet fog találni a 82 000 szóból. Az eljárásnak ezekből a jelöltekből kell kiválasztania a legmegfelelőbbet, amihez felhasználjuk a fonetikus átíratot is. A feldolgozás szavanként folytatódik, míg a teljes mondat minden szavára nem kaptunk sok-sok jelöltet. Ezután az elemkiválasztó kiválasztja a jelöltek halmazából a legjobban illeszkedőnek vélt szavakat, megkeresi a nekik megfelelőket a beszédkorpuszban, onnan kiválasztja azokat, és összefűzi őket a megszólaltatáshoz. Hogyan működik az elemkiválasztó? Kétféle költségfüggvény (KLTS-1 és KLTS-2) határozza meg, hogy az éppen vizsgált szó mennyire felel meg a kívánt követelménynek. A költségfüggvények értéke alapján történik a végleges költség szint kiszámítása. Ha ennek értéke nulla, akkor az elem 100%-osan a legoptimálisabb, ha magas szám, akkor az elem nem illeszthető, el kell vetni. A KLTS-1 határozza meg, hogy a hangsor és a hangkörnyezet szempontjából, azaz szegmentálisan mennyire felel meg a szó. Itt a szó hangsorát kell azonosítani, továbbá azt, hogy mennyire illeszkedik az öt követő, illetve megelőző szóhoz. Mivel a kiejtés folyamatos, a szóhatárokon törekedni kell arra, hogy a spektrális illeszkedés (pl. formánsmenet) is folyamatos legyen. A szó első és utolsó hangjának illeszkedését vizsgáljuk, és az illeszkedés

költségét több szempont alapján számítjuk ki. Magas költségű például, ha a szóhatáron magánhangzók találkoznak (*dunántúli áramlások*). Az ilyen szavak magas költséget képviselnek. Nulla a költség, ha a két szó egymás mellett helyezkedik el a mondatkorpuszban, hiszen ekkor a csatlakozásuk is optimális. Ebből adódik, hogy akkor nagyon optimális a keresés, ha nem szavakat, hanem szófüzéreket találunk a mondatkorpuszban. Az esetek nagy részében (ha a mondatkorpusz elég nagy) ez meg is valósul, így a szintetizált szöveg hangzása közel lesz a természeteshez. A KLTS-1 költségének kiszámításához felhasználjuk többek között azt a kutatási eredményt, hogy azonos képzési helyű mássalhangzók akusztikai megvalósulása hasonló átmeneti fázisokat okoz a hozzájuk csatlakozó magánhangzóban (Olaszy 2003), továbbá a mássalhangzók képzési módjának osztályozását és a gerjesztés fajtáját (zöngéesség-zöngétlenség). A mássalhangzók képzési helyéből adódó azonos akusztikai vetületeket az 1. táblázat mutatja be.

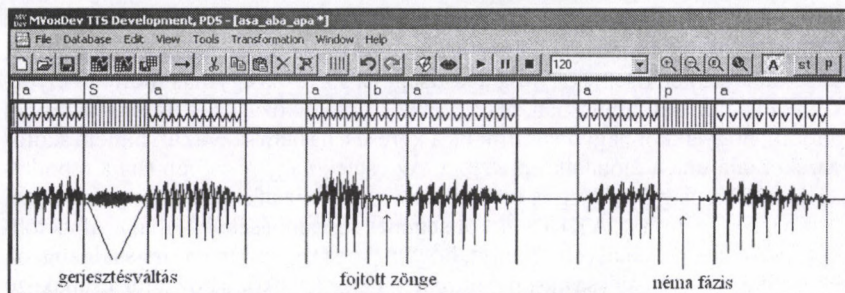
1. táblázat: A magyar mássalhangzók képzési hely és mód szerinti csoportosítása

(Az egy sorban lévő mássalhangzók hasonló akusztikai vetületet hoznak létre a hozzájuk csatlakozó magánhangzóban.)

| Képzésmód | Zárhangok | | | | | | | Zár-rés hangok | | | Részhangok | | | | | | | Nazálisok | | | | | | | |
|----------------|-----------|---|---|---|----|----|---|----------------|---|----|------------|-----|---|---|---|----|----|-----------|---|---|---|----|---|---|---|
| Képzéshely | b | p | d | t | gy | ty | g | k | c | dz | cs | dzs | v | f | z | sz | zs | s | h | m | n | ny | j | l | r |
| Két ajak | ☒ | ☒ | | | | | | | | | | | | | | | | | | ☒ | | | | | |
| Ajak-fog | | | | | | | | | | | | | ☒ | ☒ | | | | | | | | | | | |
| Fog-fogmeder | | | ☒ | ☒ | | | | | ☒ | ☒ | | | | | ☒ | ☒ | | | | ☒ | | | | ☒ | ☒ |
| Fogmeder | | | | | | | | | | | ☒ | ☒ | | | | | ☒ | ☒ | | | | | | | |
| Kemény szájpád | | | | | ☒ | ☒ | | | | | | | | | | | | | | | ☒ | ☒ | | | |
| Lágy szájpád | | | | | | | ☒ | ☒ | | | | | | | | | | | | | | | | | |
| Gége | | | | | | | | | | | | | | | | | | ☒ | | | | | | | |

Melyek tehát az optimális összefüzési pontok szóhatárok esetében? Ezt első sorban az 1. táblázat szerinti 7 artikulációs vetületi sor, illetve a beszédjel nagysága dönti el. Nem célszerű összeillesztést végezni nagy energiájú jelszakaszban (például magánhangzóban), a kis energiájú helyeket kell előnyben részesíteni. Szabad illeszteni a hangsor minden olyan pontján, ahol gerjesztésváltás megy végbe (tisztá zöngés szakaszt tiszta zöngétlen követ és fordítva, itt ugyanis a jelben intenzitásminimum keletkezik), továbbá a hangok belsejében lévő néma fázisokban, illetve zöngészakaszokban (3. ábra).

Ha tehát az akusztikai vetület ugyanaz, és például gerjesztésváltás van a két szó határán, akkor az összeillesztési költség értéke kicsi lesz, hiszen a spektrális folytonosság biztosított, és az illesztésnél kicsi az energia. Hasonló elvek alapján kialakítható az a fonetikai szabályrendszer, amellyel ki lehet jelölni a vágás konkrét helyét (a vágási pontot) az összeillesztendő szavakon belül. Erre mutat példát a 2. táblázat.



3. ábra

Az optimális illesztési pontok bemutatása az *asa*, *aba*, *apa* hangkapcsolódásoknál

(A függőleges vonalak a hanghatárokat jelzik, a zöngés hangok periódushatárait a „v” jelzésű vonalmarkerek jelzik a hullámforma felett.)

2. táblázat: Példa a fonetikai szabályrendszerből az alacsony költségű vágási pontok kijelölésére

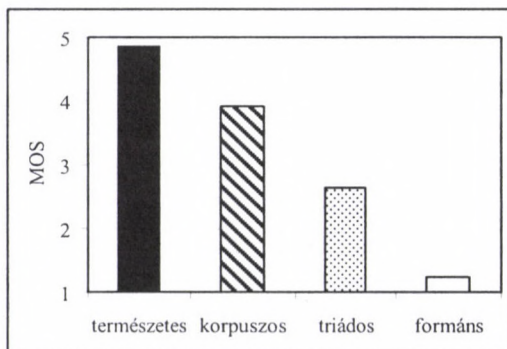
(A csatlakozó második hangot a következő szimbólumok jelölik: C = bármely más-salhangzó; V = bármely magánhangzó; $C_1 = p, t, k, ty, h, f, s, sz, c, cs$; $C_2 = v, j, l, r$; $C_3 = m, n, ny$. A hangokat a betűjelükkel adjuk meg.)

| A megelőző hang | A következő, kapcsolódó hang | A vágási pont kijelölésének szabálya | Szöveges példa (a csatlakozó hangok félkövér kiemeléssel) |
|-----------------|---|---|---|
| V | a) V | a) A hanghatár be van jelölve, ennek ellenére nem célszerű elvágni a hanghatárnál, hanem megfelelő vágási pontot kell keresni visszafelé vagy előre a hangsorban. | <i>éjszakai esőzésre</i> |
| | b) C | b) A hanghatárnál kell vágni. | <i>nyári záporok</i> |
| | a) b, d, g, gy b) b, d, g, gy | a) A hanghatárnál kell vágni. b) A hosszú hang 70%-ánál kell elvágni, a zárpfelattanás nem lesz benne. | <i>vad vihar, nagy meleg</i> <i>vad dörrenés</i> |
| c) p, t, k, ty | c) C ₁ , kivéve d) d) önmagával csatlakozik | c) A hanghatárnál kell vágni. d) A hosszú hang 70%-ánál kell elvágni, a zárpfelattanás nem lesz benne. | <i>szép sereg</i> <i>sok kis</i> |
| | e) V, C ₂ , C ₃ | e) A hanghatárnál kell vágni. | <i>szép felhők</i> |
| | a) C, kivéve m b) önmagával | a) A hanghatárnál kell vágni. b) A hang 70%-ánál kell elvágni. | <i>nem volt</i> <i>nem marad</i> |

A KLTS-2 határozza meg, hogy a hangsor és hangkörnyezet szempontjából kiválasztott szó, szófüzér mennyire felel meg a prozódiai követelményeknek. Itt szempont az is, hogy a kiválasztott szó a mondatkorpusz ugyanazon mondatában szerepel-e, mint az előző. Ha igen, akkor a költséget ez a tény is csökkenti. A prozódiai költség meghatározásánál – az időtengelyi pozíción felül – felhasználjuk az F_0 értékének a változását is. Ha nagy F_0 -ugrás van a két szó között, akkor a költség magas lesz, tehát a két elem nem illeszthető össze.

A működő rendszer minősítése

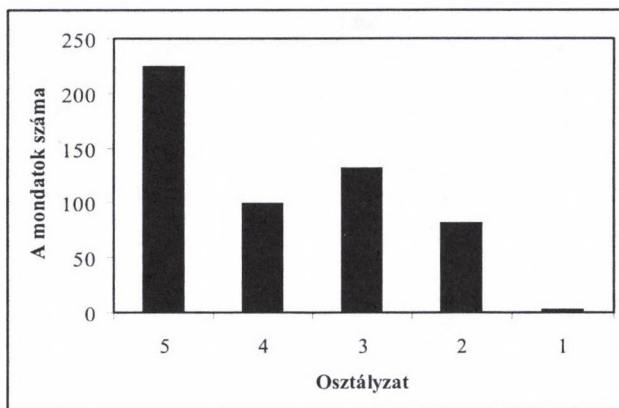
A kísérleti, időjárás-jelentést felolvasó, korpuszalapú beszéd szintetizáló rendszer minőségét percepciós teszttel vizsgáltuk. Három magyar rendszert hasonlítottunk össze, a Multivox formánsszintetizátort, a Profivox diád-triád elemösszefűzéses rendszert és a korpuszos felolvasót. A tesztelőktől a hangminőségre vonatkozó ítéletet kértünk 5 pontos skála szerint: nagyon jó (5), jó (4), átlagos (3), gyenge (2), elfogadhatatlan (1). A teszt anyagát a webről kiválasztott időjárás-jelentés 10 mondata alkotta. Ezeket állítottuk elő a fenti rendszerekkel, valamint felolvastattuk őket a beszédkorpusz eredeti bemondójával is. A tesztelőknek tehát összesen 40 mondatot kellett meghallgatni véletlen sorrendben. Minden mondatra egy ítélet született. A tesztet egy interaktív honlap segítségével bonyolítottuk le. A mondatokat 221 személy (egyetemi hallgatók, 185 férfi és 36 nő) hallgatta meg. A teszt elején ismerkedésképpen minden mondattípusból egy-egy mondatot meghallgathattak. A tesztben a mondatokat csak egyszer hallották, ismételésre nem volt mód. A meghallgatások csendes, otthoni környezetben, átlagos (nem professzionális) hangszórókon, illetve fejhallgatókon történtek (a tesztelők a meghallgatás körülményeire vonatkozóan is kitöltöttek adatokat a teszt előtt). A teszt eredményeit a 4. ábrán mutatjuk be. A tesztelés eredményét a MOS (Mean Opinion Score) értéke fejezi ki.



4. ábra

A szubjektív minősítés átlagai az egyes szintézisteknológiákra

Az értékelésekből látható, hogy a korpuszalapú szintézis hangminősége magasban kiemelkedik a másik két technológiával szemben. A szó-, szófüzéralapú összeillesztéssel tehát átléptünk a percepció megítélésben egy olyan határt, amelyet a hullámforma-összefűzéses rendszereknél még nem tudtunk elérni, annak ellenére, hogy ott is emberi beszéd részleteit fűztük össze. Feltételezzük, hogy a szó képviseli azt a mondatépítő elemet, amelynek szintjén már elégséges egyéni hangjellegzetesség van jelen a hullámformában, hogy a hallgató a beszélő hangszínezetét, egyéni stílusát is felismerje, és ennek folytán értékítéletével megközelítse a jó (4) szintet. Természetesen a korpuszalapon szintetizált mondatokban is vannak egyenetlenségek a hullámforma folytonosságát illetően (dal- lamugrások, hangszínezet-változások stb.), de úgy tűnik, ha kevés van ezekből, akkor az összegzett ítéletek meghozásakor ezeket a percepció mechanizmusunk ugyanúgy tűri, feldolgozza, mint az olvasásnál a felolvasási mechanizmusunk a betűkimaradásokat, betűhibákat. Elvégeztünk egy másik percepció értékelést is, amelyben a rendszer belső működését, a mondat-összeállítás hatásosságát vizsgáltuk. Kíváncsiak voltunk, hogy nagyszámú időjárás-jelentés mondat előállítása esetén hogyan alakul a mondatok hangminőségi eloszlása. A teszthez 540 mondatot szintetizáltunk (olyanokat, amelyek nem szerepelnek a rendszer korpuszában), és ezek hangminőségét értékeltettük 4 fővel (30–60 éves férfiak). A mondatokat a tesztelők fejhallgatón hallgatták. Egy tesztelő egy alkalommal 135 mondatot hallgatott meg. Egy mondatot többször is meghallgathattak. Az értékelést az előző teszthez hasonlóan egy 1–5 osztályzatú skálán kértük, a következők szerint: 5 = nagyon jól érthető, mintha bemondó olvasta volna; 4 = jól érthető; 3 = közepesen érthető; 2 = nehezen érthető, 1 = nem érthető. Az eredményeket az 5. ábrán mutatjuk be.



5. ábra

A mondatok hangminőségi eloszlása 540 időjárás-jelentés mondaton mérve

A tesztelők 225 mondatra adtak nagyon jó értékelést, 99-re jót és 132-re közepesen érthetőt. Az összes mondatra számolva ez 90%-os lefedést jelent. Mindössze 81 mondatnál volt gondjuk a minőséggel, itt nehezen érthető ítéletet adtak. 3 mondat esetében nem lehetett megérteni a mondat tartalmát. Ezek az eredmények azt mutatják, hogy a jelen korpuszalapú szintetizáló rendszernél az esetek többségében igen jó hangminőség érhető el. Előfordulhatnak azonban igen kis számban olyan generált mondatok, amelyeknél a költségfüggvény gyakorlatilag abszolút rossz döntéseket hoz. Ilyen mondat volt például a következő: *Vasárnap számmottevő eső csak Északkelet-Magyarországon esik, hajnalban és délelőtt.* Mindezekből az is látszik, hogy a meggyőzően jó MOS eredmények mellett számolni kell azzal is, hogy rossz válogatás esetén a korpuszalapú szintetizálás is adhat igen rossz minőségű beszédet. A jelen kísérleti rendszerben ez csak 0,6%-ban fordult elő. A hibák eredhetnek abból, hogy a mondatkorpusz és a neki teljes mértékben megfeleltetett beszédkorpusz között nem teljes az egyezés, azaz hanghatár, illetve átírási hibák vannak. Továbbá a prozódiai modell döntései is hibásak lehetnek. A hibák kijavítására a konkrét hibakeresésen és hibajavításon túl az egyik megoldás az lehet, hogy az eredeti bemondóval felolvastatjuk a rossznak ítélt mondatokat, és hozzákapcsoljuk a korpuszhoz, kihasználva, hogy ez egy nyitott beszédkorpusz.

Összefoglalás

A tanulmányban bemutattuk az első magyar nyelvű, korpuszos technológián alapuló szövegszintetizáló kísérleti változatát. A rendszerben alkalmazott módszer gyökeresen eltér a korábbi beszéd-szintézis-technológiákra jellemző módszerektől. A rendszert kötött témakörre készítettük el, időjárás-jelentéseket tud felolvasni. Az eddigi teszteredmények azt mutatják, hogy ez a legújabb technológia igen jó minőségű beszédet biztosíthat: az esetek nagy részében nem lehet megkülönböztetni, hogy szintetizált-e a mondat, vagy egy bemondó olvasta-e fel.

Irodalom

- Kawai, Hisashi – Toda, Tomoki – Ni, Jinfu – Minoru, Tsuzaki – Tokuda, Keiichi 2004. Ximera: a new TTS from ATR based on corpus-based technologies. In: *Proceedings of the 5th ISCA Speech Synthesis Workshop (SSW5)*. Pittsburg, 642–645.
- Mihajlik, Péter – Révész, Tibor – Tatai, Péter 2002. Phonetic transcription in automatic speech recognition. *Acta Linguistica Hungarica* 49/3–4. 407–425.
- Nagy András – Pesti Péter – Németh Géza – Böhm Tamás 2005. Korpusz-alapú beszéd-szintézis rendszerek megvalósítási kérdései. *Híradástechnika* 2005. január, 18–24.
- Olaszy, Gábor – Gordos, Géza – Németh, Géza 1992. The MULTIVOX multilingual text-to-speech converter. In: Bailly, G. – Benoit, C. – Sawallis, T. (eds.): *Talking machines: Theories, models and applications*. Elsevier, Amsterdam, 385–411.
- Olaszy Gábor 1999. Beszédatadtbázisok készítése gépi beszédelőállításához. *Beszédkutatás* '99. 68–89.

- Olaszy Gábor 2003. Az artikuláció akusztikai vetülete – a hangsebészet elmélete és gyakorlata. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia a gyakorlatban*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 241–254.
- Olaszy Gábor 2006. *Hangidőtartamok és időszervezeti elemek a magyar beszédben*. Nyelvtudományi értekezések 155. Akadémiai Kiadó, Budapest.
- Olaszy, Gábor – Németh, Géza – Olaszi, Péter – Kiss, Géza – Zainkó, Csaba – Gordos, Géza 2000. Profivox – a Hungarian TTS system for telecommunications applications. *International Journal of Speech Technology* 3/3–4. 201–215.
- Schweitzer, Antje – Braunschweiler, Norbert – Klankert, Tanja – Möbius, Bernd – Sauerlich, Bettina 2003. Restricted unlimited domain synthesis. *Proceedings of Eurospeech 2003*. Geneve, 1321–1324.

A szerzők köszönetüket fejezik ki Pesti Péternek, aki a rendszer kódját programozta, és Mihajlik Péternek, aki rendelkezésükre bocsátotta a jelen munkában használt beszédfelismerési eszközöket, és segítséget nyújtott azok használatában.

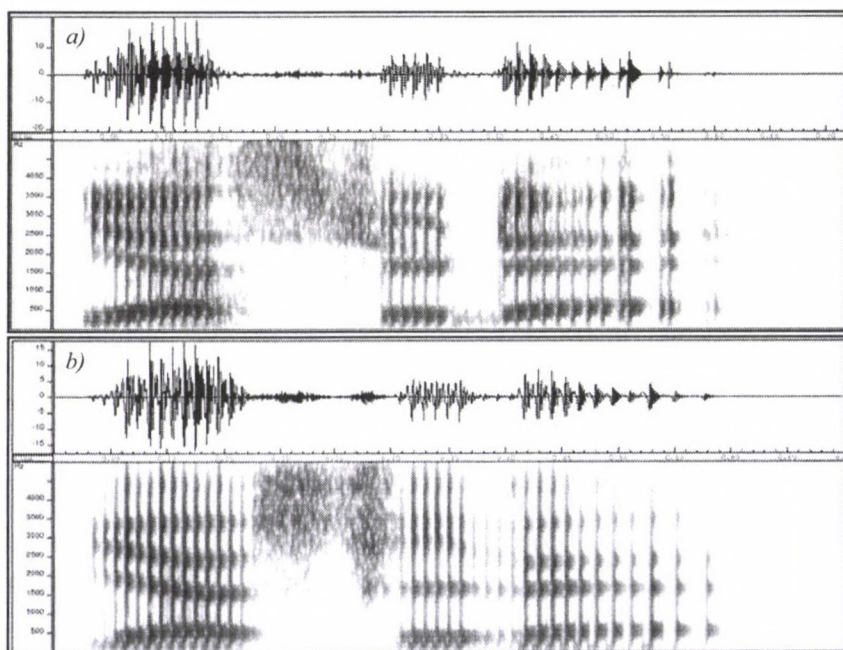
Ez a kutatás az NKFP 2/034/2004 szerződése alapján támogatásban részesült.

A GLOTTALIZÁCIÓ SZEREPE A BESZÉLŐ SZEMÉLY FELISMERÉSÉBEN

Bóhm Tamás

Bevezetés

A magánhangzók és zöngés mássalhangzók artikulációjának alapja a zönges (Gósy 2004: 28). A zöngékepzés, a fonáció során a hangszalagok általában közel periodikusan (kváziperiodikusan) rezegnek. Ilyenkor a hangszalagok nagyjából állandó időközönként összecsapódnak – a rezgés reguláris. Rövidebb-hosszabb ideig azonban ez a rezgés lehet irreguláris. Ekkor a hangszalagok összecsapódásai között eltelt idő széles határok között ingadozik, és általában jóval hosszabb. Ezt a jelenséget nevezzük glottalizációnak (1. *a* ábra).



1. ábra

A *yesterday* szó férfiejtésben, glottalizált végződéssel:

a) természetes beszéd, *b*) másolatszintézis (1. a hanganyag ismertetésénél)

A glottalizációt a szakirodalomban többféleképpen definiálták (Laver 1980: 93–95; Huber 1992; Slifka 2000: 100–101; Redi–Shattuck–Hufnagel 2001). Ebben a tanulmányban az „érzékeltetőn irreguláris hangszalagrezgést” tekintjük glottalizációnak. Az irregularitás jelentkezhethet a rezgés pillanatnyi frekvenciájában, amplitúdójában vagy mindkettőben. Az ingadozás mértékének elég nagynek kell lennie ahhoz, hogy a hallgató számára érzékeltetően eltérjen a reguláris zöng hangzásától. Ez a kikötés azért fontos, mert reguláris zöngképzés esetén is csak kváziperiodikus rezgésről beszélhetünk, amelynek frekvenciája és amplitúdója kismértékben ingadozik (Gordos–Takács 1983: 24).

A glottalizációt – a fenti definíció szerint – akusztikai jelenségnek tekintjük, nem feltételezünk semmit a produkciójáról vagy percepciójáról (az említett érzékeltetőségen felül). Ettől függetlenül természetesen vannak produkciós és percepciós vonatkozásai is, ezeket a következőkben foglaljuk össze.

A glottalizáció produkcióját hagyományosan a hangszalagok szoros összeszorításával magyarázzák (Laver 1980: 122–126), ami a reguláris fonációnál erősebb záródást eredményez. Ezért a rezgés kevésbé stabil, frekvenciája és amplitúdója ingadozhat. Slifka (2000) azonban a tüdő nyomásszintjének közvetett mérésével kimutatta, hogy glottalizációt a hangszalagok összeszorításával és ellazításával is el lehet érni.

A glottalizációt érdes, recsegő hangként érzékeljük. Az angol nyelvű szakirodalomban a hozzá tartozó érzeti jelenséget, a hangminőséget számos névvel illetik (*creaky*, *vocal fry*, *pressed* stb.). Catford megfogalmazásában a glottalizáció ahhoz hasonlít, amikor valaki egy rudat végighúz egy fémkerítésen (idézi Laver 1980: 122). Egy magyar tanulmány is kiemeli, hogy a szélsőségesen mélyen futó alaphfrekvencia-görbét nem a hang mélyüléseként, hanem nyikorgóvá, torzzá válásaként érzékeljük (Markó 2005: 61–62).

Több angol nyelvre és egy magyar nyelvre végzett vizsgálat is nagy különbségeket mutatott ki az egyes beszélők fonációs szokásaiban. Ezekből a tanulmányokból arra következtethetünk, hogy a glottalizáció gyakorisága jellemző a beszélő személyre. Redi és Shattuck–Hufnagel (2001) 14 amerikai angol anyanyelvű bemondó beszédét elemezte. Volt olyan adatközlő, aki a vizsgált esetek 88%-ában, míg egy másik személy csak 13%-ában glottalizált. A négy amerikai személy, akik Slifka (2000: 100–103) vizsgálatában szerepeltek, a bemondások 5–93%-ának a végén glottalizált. Valószínűleg magyar beszélők között is hasonló különbségek vannak: Markó (2005: 61) egy adatközlő spontán beszédében gyakran, a többi három felvételeiben csak néhányszor észlelt „nyikorgó zöngét”.

A glottalizáció az amerikai angolban gyakori jelenség intonációs frázisok elején (Redi–Shattuck–Hufnagel 2001), de több tanulmány szerint legnagyobb arányban a bemondások végén jelenik meg (Henton–Bladon 1987; Redi–Shattuck–Hufnagel 2001). Ez utóbbi kiemeli, hogy a beszélők között ezen a pozíción a legnagyobb különbség. Slifka (2000) ezzel összhangban az egyes személyek közötti különbségeket szintén a bemondások végén mutatta ki.

Összefoglalva, a glottalizáció előfordulási aránya jellemző a beszélő személyére (elsősorban a bemondások végén). Ennek háttérében valószínűleg fiziológiai, szociolingvisztikai és egyéb hatások állnak.

Hipotézisek és a kísérlet alapelve

Kizárólag a hangjuk alapján az emberek képesek ismerős személyeket felismerni. Ebben a folyamatban számos egyéni ismertetőjegy játszhat szerepet. Ezek a beszéd olyan akusztikai jellemzői közül kerülhetnek ki, amelyek az egyes bemondókra következetesen eltérőek (Laver 1980: 9–10). Beszélőnként változik, hogy a bemondások vége milyen gyakran glottalizált (gyakran, ritkán, vagy néha igen, néha nem). Így ez olyan egyéni tulajdonság, amely segítheti a hallgatót a beszélő felismerésében. Az viszont nyitott kérdés, hogy ennek a jelenségnek tényleg van-e szerepe a felismerési folyamatban, azaz megjegyezzük-e ismerős személyek jellemző fonációs szokásait. Ezt a kérdést az alábbi két hipotézisben pontosítottuk:

1. A végső glottalizáció érzékelhető a hallgatók számára.
2. A végső glottalizáció szerepet játszik az ismert beszélő személy felismerésében – egyes beszélők és egyes hallgatók esetén.

Az első hipotézist két típusú hangz óanyagra már bizonyították: szintetikus kiktartott magánhangzókra és természetesen szópárookra (Huber 1992). A jelen tanulmány egyik célja, hogy egy-két szavas szintetikus beszéd alkalmazásával közvetlenebb bizonyítékot nyerjen erre a hipotézisre. Grassegger (2003) megállapításai és személyes tapasztalatok alapján a hipotézis igazolása nagyon valószínű.

A második hipotézist nem jelenthetjük ki minden beszélőre, mert a korábbi tanulmányok szerint nem minden bemondónál jelentkeztek markáns egyéni különbségek. Minden hallgatóra se általánosíthatunk, mert egyrészt az egyes hallgatók által a felismerésre használt ismertetőjegyek eltérőek lehetnek (van Lancker et al. 1985), másrészt az egyes hallgatók különbözőképpen érzékelik az egyes fonációs típusokat (Kreiman et al. 1992).

Percepciós kísérlet segítségével próbáltuk igazolni a hipotéziseket, formánszintézis-technológia alkalmazásával. A módszer hasonlít Allen és Miller (2004) kísérletéhez. Négy beszélő 4, egy-két szavas felvétele alapján elkészült azok szintetikus másolata. Ezekből a másolatokból két változat született: az egyik végén a hangszalagrezgés reguláris, a másik végén irreguláris, azaz glottalizált. Az eredeti bemondások végén két bemondó megbízhatóan glottalizált, kettő pedig ritkán, ezért a két mesterséges változat közül az egyik megfelelt a beszélő szokásos végső hangjellegzetességének, a másik nem. Ezek a párok lehetőséget adtak annak vizsgálatára, hogy a szokásos (glottalizált vagy reguláris) végződéstől való eltérés milyen hatással van a beszélő személy azonosítására.

A végső glottalizáció mellett az alaphfrekvencia-görbét is módosítottuk (ennek módját később ismertetjük). Egy felvétel alaphfrekvencia-görbéje alapján bizonyos pontossággal megállapítható többek között az F_0 átlaga, amely bizonyítottan a beszélő személy robusztus ismertetőjegye (irodalmi áttekintés: Pardo-Remez 2005). Így ha egy felvételre egy másik beszélő alaphfrekvencia-görbéjét

ültetjük át, akkor – tapasztalataink szerint – a beszélő személye nehezen felismerhető.

Ezzel a drasztikus módosítással több célt is elértünk. Ha mindig ugyanaz a különbség a párok két tagja között (glottalizált – nem glottalizált), akkor a kísérleti személyek hamar rájöhetnek a hipotézisekre és a módszerekre, így a továbbiakban a beszélőfelismerési folyamattól független stratégiát követhetnek. Ezenfelül az egymástól csak az alapfrekvencia-görbében eltérő párok kontrollként szolgáltak: egyrészt igazolták a kísérleti módszer működését, másrészt összehasonlítási alapot jelentettek. Az F_0 -átlag robusztusságából adódóan az alapfrekvencia-görbére számolt eredmény a glottalizációra kapott eredmény felső korlátja. Így azt feltételeztük, hogy a glottalizáció hatása az alapfrekvencia-görbe hatásánál jóval kisebb, de kimutatható.

Így mindkét szintetikus változatból (glottalizált és nem glottalizált) további két változat született: egy az eredeti és egy a módosított alapfrekvencia-görbével.

A kísérleti személyek minden lehetséges változatpárt meghallgattak, és eldöntötték, hogy melyik hasonlít jobban a megadott személy hangjára. Az összes kísérleti személy jól ismerte mind a négy beszélőt. Ha a kísérleti személyek gyakran felismerik a beszélő szokásos végső hangminőségét, akkor az alátámasztja a hipotéziseket. Ha a válaszok véletlenszerűek, akkor a hipotézisek valószínűleg tévesek, vagy a kísérlet nem elég érzékeny. A kísérletet amerikai angol nyelvre végeztük, mert a fentebb hivatkozott cikkek többsége ezzel a nyelvvvel dolgozott, valamint a munka jelentős részét a szerző egy amerikai egyetem vendéghallgatójaként végezte el.

Anyag, módszer, kísérleti személyek

Egy saját készítésű korpusz 9 amerikai angol bemondója közül választottuk ki a kísérletben részt vevő 4 személyt. Minden adatközlőre kiszámítottuk a glottalizációs arányt (a végső pozíciók hány százalékában glottalizált). Egyes adatközlők szinte mindig glottalizáltak, mások ritkán, és megint mások néha igen, néha nem. A kísérlethez két gyakran és két ritkán glottalizáló személyt választottunk (1. táblázat) – ha igaz a második hipotézis, akkor egy ilyen éles ellentét esetén valószínűleg kimutatható. A két férfi (FR és FG) átlagos alapfrekvenciája jelentős eltérést mutatott. Ez a kísérlet eredményeit nem befolyásolja, mert a hipotézisek vizsgálata során a kísérleti személyek ugyanahhoz a beszélőhöz tartozó, kétféle módon szintetizált hanganyag közül választották ki a beszélőre jellemzőbbet (nem két beszélőt hasonlítottak össze).

A négy adatközlő négy-négy rövid (egy-két szavas) olvasott beszédét használtuk. Az összesen 16 felvételtől másolatszintézist készítettünk egy Klatt-típusú formánsszintetizátorral (Klatt–Klatt 1990). A másolatszintetizálás olyan eljárás, amely során olyan szintézisparamétereket keresünk, amelyekkel az eredeti felvételre minél jobban hasonlító hullámformát kapunk (1. b) ábra). Az így előállított másolatokon mindig megjelenik az eredeti felvételnek megfelelő glottalizáció is.

1. táblázat: A bemondók adatai

| Azonosító | Nem | Végső glottalizáció | Átlagos F_0 |
|-----------|-------|---------------------|---------------|
| NR | nő | ritkán (20%) | 232 Hz |
| FR | férfi | ritkán (9%) | 136 Hz |
| NG | nő | gyakran (93%) | 210 Hz |
| FG | férfi | gyakran (83%) | 94 Hz |

Minden egyes másolatból a végső glottalizáció és az alapfrekvencia-görbe módosításával több változatot állítottunk elő. Az utóbbi módosítás esetén a másik, azonos nemű beszélő stilizált alapfrekvencia-görbéjét használtuk a szintézishez. A görbét időben úgy vetemítettük, hogy a csúcsok ugyanazokra az időpontokra essenek, mint az eredeti görbe csúcsai. Így ez a szintetikus változat az alapfrekvencia szempontjából (pl. átlagos F_0) egy másik személyre hasonlított, de minden más akusztikai tulajdonság (pl. beszédtempó, formánsok) változatlan maradt.

A kétféle módosítás alkalmazásával mind a 16 eredeti felvételhez négy szintetikus változat készült (összesen 64 darab): 1. másolatszintézis: minden szempontból hasonlít az eredetire; 2. módosított glottalizáció; 3. módosított alapfrekvencia-görbe; 4. módosított glottalizáció és alapfrekvencia-görbe.

Kilenc kísérleti személy (5 nő és 4 férfi) vett részt a percepciókísérletben. Mindegyikük vagy angol anyanyelvű volt (5 fő), vagy legalább három éve folyamatosan angol nyelvtérületen élt (4 fő, anyanyelvük héber, japán, koreai és mandarin). Életkoruk 23–29 év volt, és nem ismerték a kísérlet hipotéziseit és módszereit. Feltételezhetően jól ismerték a négy adatközlő hangját, mert legalább egy éve ugyanabban a kutatócsoportban dolgoztak, vagy évfolyamtársaik voltak.

A kísérletet egy Matlab 7.0-ban írt grafikus program vezérelte. Ez a program játszotta le a hangzó anyagot, jelenítette meg a kérdéseket, és rögzítette a válaszokat. A program mindig minden válaszlehetőséghez egy külön nyomógombot jelenített meg. A válaszokat utólag nem lehetett módosítani, és minden stimulus csak egyszer hangzott el (újrajátszásra nem volt lehetőség). A hanganyagokat a program minden kísérleti személynek külön, véletlenszerű sorrendbe állította.

A kísérlet négy részből állt. Az első ellenőrizte, hogy a hallgatók tényleg jól ismerik-e a beszélők hangját, a második és a harmadik a szintetikus hanganyag minőségét tesztelte, míg a negyedik rész a hipotéziseket vizsgálta. Az első részben csak az eredeti, természetes felvételeket használtuk, míg a többiben csak szintetikusakat.

1. rész. Olyan kísérleti személyekre volt szükségünk, akik jól ismerik a beszélők hangját. Annak érdekében, hogy ezt biztosítsuk, és esetleg kiszűrjük az alkalmatlan hallgatókat, a kísérletet egy beszélőfelismerési feladattal kezdtük. A feladat során egy egy-két szavas felvétel meghallgatása után egy listáról ki kellett választani a hallott beszélő nevét. A négy ismert beszélő mellett két olyan személy felvételeit is lejátszottuk, akiknek a hangját nem ismerték a kísérleti

személyek. Ennek megfelelően a nevek listáján is szerepelt két további tétel „más nő” és „más férfi” felirattal.

2. rész. A cél a formánsszintézissel előállított glottalizált beszéd percepciójának mérése volt: vajon hasonló érzetet vált-e ki a hallgatókból a természetes és a mesterséges glottalizáció. A feladat páronkénti összehasonlítás volt. A párok két tagja mindig ugyanattól a beszélőtől származó, ugyanazt a szöveget realizáló szintetikus hanganyag volt, de az egyik vége glottalizált volt, a másiké nem. A két beszédrészlet lejátszása után a hallgatónak válaszolnia kellett arra a kérdésre, hogy „Melyik felvétel érdekesebb hangzású?” A választ egy 6 pontból álló skálán adhatta meg, ahol a két szélső érték mellett a „határozottan az első” és a „határozottan a második” felirat jelent meg. Ennek a skálának az előnye, hogy a hallgató egyetlen kattintással megválaszolja az eldöntendő kérdést, és egy bizonyosságértéket is megad (3, 4: kicsi; 1, 6: nagy).

3. rész. Ebben a részben azt ellenőriztük, hogy felismerhető-e a beszélő személye a szintetikus hanganyag alapján. A feladat ugyanaz volt, mint az első részben, csak itt nem a természetes, hanem a szintetikus bemondásokat játszottuk le, és nem használtuk az ismeretlen személyek felvételeit.

4. rész. Ennek a résznek a célja a hipotézisek vizsgálatára volt. A feladat páronkénti összehasonlítás volt, ahol a párok mindkét tagja ugyanannak az eredeti felvételnek két szintetikus változata volt. Így a beszélő és a nyelvi tartalom ugyanaz, de vagy a glottalizáció, vagy az alapfrekvencia-görbe, vagy mindkettő szempontjából eltér a két felvétel. A képernyőn megjelenő kérdés magyar fordításban: „Melyik XY hangja (vagy melyik áll közelebb hozzá)?” – ahol XY a beszélő személy neve. A válaszokat ismét a hatelemű skálán kellett megadni, mint a második részben. A felvételek 4 szintetikus változatából minden lehetséges párosítást felhasználunk, mindkét sorrendben.

Az adatok elemzését az SPSS 14.0-ban elérhető leíró statisztikák, egymintás és kétmintás t -próba, valamint ANOVA segítségével végeztük. Minden adatot 3 tizedesjegy pontossággal jelentettünk meg. Így ha a program számításai szerint $p = 0,000$, akkor azt $p < 0,0005$ -ként értelmezzük és közöljük.

Eredmények

1. rész. Beszélőfelismerés természetes hanganyag alapján

A válaszok 68%-a megegyezett a tényleges beszélő személyével. Amint a módszer leírásánál említettük, ebben a részben két, a hallgatók számára ismeretlen személy hangját is felhasználtuk. A helyes azonosítások aránya a négy ismert személyre 75% volt, a két ismeretlenre pedig jóval alacsonyabb, 53%. Az egyes kísérleti személyekre 54% és 79% közötti felismerési arányokat kaptunk az ismert és ismeretlen beszélőkre együttesen. Eszerint – ahogy az várható volt – nem minden személy ismerte azonos mértékben a beszélőket. Az alábbiakban belátjuk, hogy mindegyikük esetén elvethetjük azt a nullhipotézist, hogy számukra ismeretlenek voltak a beszélők.

Ha egy kísérleti személy nem tudta volna felismerni a beszélőket, akkor válszai véletlenszerűen oszlottak volna el a lehetséges válaszelehetőségek között. Így 1/3 körüli felismerési arányt érhetett volna el, mert minden kérdésnél 3 férfi vagy 3 női válaszgomb közül választott volna találgatással (feltételezhetjük, hogy a beszélő nemét meg tudta állapítani). Minden hallgatóra ennél szignifikánsan magasabb értéket kaptunk (egymintás t -próbával minden esetben $t \geq 2,912$ és $p \leq 0,005$), ami kizárja a találgatás lehetőségét. Eszerint minden hallgató számára ismerős volt a négy adatközlő hangja.

A felismerési arány a két ismert férfi esetén sokkal magasabb volt (89% és 100%), mint a két ismert nőnél (54% és 57%). Ennek oka valószínűleg a két férfigang közötti nagy átlagos F_0 -különbség (1. táblázat).

2. rész. Mesterséges glottalizáció és az érdekesség

Az adatok elemzésénél azt vizsgáltuk, hogy vajon a hallgatók a párnak azt a tagját tekintették-e érdekesebbnek, amelyik végét glottalizáltként szintetizáltuk. Az eredmények azt mutatták, hogy 66%-ban igen (standard hiba: 3%), ami szignifikánsan eltér az 50%-os véletlenszerű átlagtól ($t = 5,711$; $p < 0,0005$). Tehát a szintetikus glottalizáció felhasználható a természetes helyett.

3. rész. Beszélőfelismerés szintetikus hanganyag alapján

Szintetikus hanganyag esetén a beszélők felismerési aránya 68% volt, ami nem tér el szignifikánsan az első részben, természetes hanganyagokra számolt átlagtól (kétmintás t -próbával az összes beszélőre $t = 0,157$ és $p = 0,876$; az ismerősökre $t = 1,938$ és $p = 0,053$). Ezért a kísérletben alkalmazhatjuk a szintetikus hangzó anyagot a hipotézisek vizsgálatára.

4. rész. A glottalizáció hatása a beszélőfelismerésre

A felvételek négy szintetikus változatát összesen hatféleképpen párosítottuk össze (2. táblázat). A hanganyagok sorrendje a táblázatban nem lejátszási sorrendet jelent, mert a párokat mindkét sorrendben felhasználtunk. A 6-értékű válaszokat egy kétértékű döntésre (első vagy második hanganyag) és egy háromértékű bizonyosságra választottuk szét. A hipotézisek vizsgálatát a kétértékű döntéseken végeztük.

Egy választ akkor tekintettünk „találatnak”, ha a kísérleti személy a beszélőre jellemzőbb – azaz a 2. táblázat „Egyik hanganyag” oszlopában szereplő – változatot választotta. Ezek az alaphékvencia, a glottalizáció vagy mindkettő szempontjából az eredeti felvételre hasonlítanak.

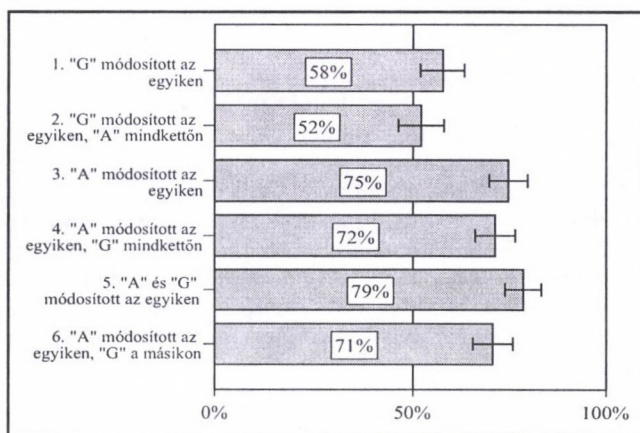
A hat pártípushoz tartozó találati arány (a válaszok hány százaléka találat) a 2. ábrán látható. A hipotézisek szempontjából az első pártípus az érdekes. Ha a végső glottalizáció léte vagy nem léte nem jelent érzékelhető különbséget, vagy ha ez a jelenség nem játszik szerepet a beszélő személy azonosításában, akkor véletlenszerű válaszokat kapnánk (találgatás). Ez azt jelenti, hogy a találati arány körülbelül 1/2 lenne, mert a kísérleti személyek kétértékű döntéseket hoztak. Ha azonban a hipotézisek igazak, akkor az arány ennél magasabb kell, hogy legyen. A mért találati arány 58%. Ez az arány nem sokkal magasabb 50%-nál, de az eltérés szignifikáns (egymintás t -próbával $t = 2,619$; $p = 0,009$). A kis elté-

rés azt jelzi, hogy a glottalizáció megváltoztatásának nincs olyan alapvető hatása a beszélőfelismerésre, mint az alapfrekvenciának. Az alacsony p érték viszont arra utal, hogy – bár kismértékű és nem általános érvényű – létezik ez a hatás, azaz a hipotézisek igazak.

2. táblázat: A hat pártípus

(E = eredeti, M = módosított, a beszélőre nem jellemző.)

| | Egyik hanganyag | | Másik hanganyag | |
|---|-----------------|----------------------|-----------------|----------------------|
| | Glottalizáció | Alapfrekvencia-görbe | Glottalizáció | Alapfrekvencia-görbe |
| 1 | E | E | M | E |
| 2 | E | M | M | M |
| 3 | E | E | E | M |
| 4 | M | E | M | M |
| 5 | E | E | M | M |
| 6 | M | E | E | M |



2. ábra

Találati arány a pártípusokra

(G = glottalizáció, A = alapfrekvencia-görbe;

az ábrán a 95%-os konfidenciaintervallumokat is feltüntettük.)

A találat változóra ANOVA-t végeztünk, ahol a hallgató véletlen faktor, a beszélő, a bemondás szövege és a pártípus rögzített faktor volt. A faktorok értékeinek páronkénti összehasonlítását Tukey-féle post hoc próbákkal végeztük 95%-os szinten.

A pártípus faktor szignifikáns volt ($F = 18,327$; $p < 0,0005$), és a post hoc próbák két homogén csoportra bontották értékeit (amelyek tagjai között nem

mutatható ki szignifikáns eltérés, viszont a csoportok között igen): az egyikbe az 1. és 2. pártípus, a másikba a 3–6. pártípusok tartoztak. Ahogy az ANOVA és a 2. ábra is mutatja, az első pártípus találati aránya alacsonyabb a harmadik pártípusénál, ahol az eredeti és módosított alapfrekvencia-görbéjű változatot hasonlítottuk össze. Ez azt jelzi, hogy feltevéseinknek megfelelően a beszélőfelismerési folyamatban az alapfrekvencia-görbe által hordozott információ nagyobb súlyú, mint a végső glottalizáció léte vagy nem léte.

Bár az ANOVA nem mutatott ki különbséget az első és a második pártípus találati aránya között, az első szignifikánsan eltér az 50%-tól (lásd fent), a második pedig nem (egymintás t -próba: $t = 0,706$; $p = 0,480$). Azaz ha mindkét felvételen módosítjuk az alapfrekvencia-görbét, akkor valószínűleg már egyik felvétel se hasonlít az eredeti beszélőre, így a válaszadók csak találgatni tudtak.

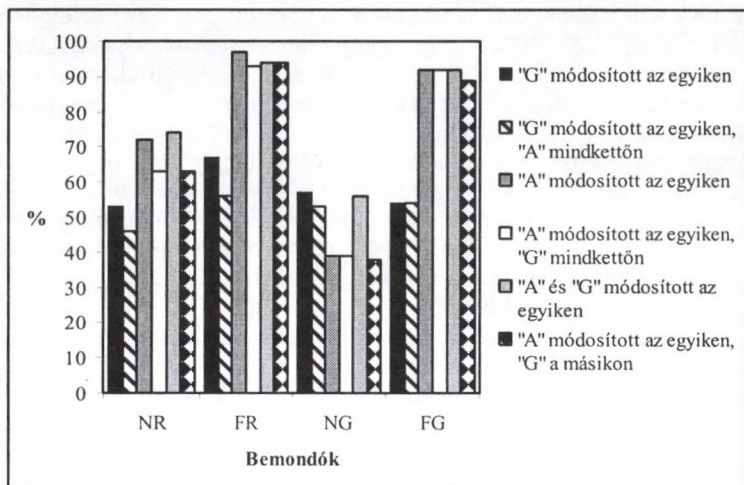
A negyedik pártípusnál is enyhén csökkent a felismerési arány a harmadikhoz képest (75-ről 72%-ra). A legmagasabb arányt akkor kaptuk, amikor a hallgatók az eredeti felvételhez hasonló szintetikus mintát hasonlították össze azzal, amelynek mindkét vizsgált tulajdonságát módosítottuk (ötödik pártípus, 79%). A 3–6. pártípusokból álló homogén csoportban akkor kaptuk a legalacsonyabb találati arányt, amikor az egyik felvétel a végső glottalizáció, a másik az alapfrekvencia-menet szempontjából hasonlított az eredetire (hatodik pártípus, 71%). Ezek a kis, nem szignifikáns különbségek a glottalizáció hatásával magyarázhatók.

A pártípuson kívül a beszélő faktor volt szignifikáns ($F = 19,730$; $p < 0,0005$). A két férfi beszélő találati aránya (79% és 84%) magasabb volt, mint a két nőé (47% és 62%), hasonlóan a kísérlet harmadik részében tapasztaltakhoz. A jelenség hátterében valószínűleg az eltérő szintézisminőség áll.

A kísérleti személy faktor különböző szintjeihez tartozó találati arányok között nem volt szignifikáns különbség ($F = 1,150$; $p = 0,372$). Eszerint nem tudunk a hallgatók egyéni jellegzetességeivel vagy eltérő anyanyelvével összefüggésbe hozható különbséget kimutatni az eredményekben.

Az egymásra hatások közül kettő volt szignifikáns: a hallgató-beszélő ($F = 3,576$; $p < 0,0005$) és a pártípus-beszélő ($F = 6,156$; $p < 0,0005$). Az előbbi szerint a beszélőfelismerési folyamat függ a beszélőtől és a hallgatótól is, azaz indokoltan korlátoztuk a második hipotézist egyes hallgatókra és egyes beszélőkre. A pártípus-beszélő párosítások közötti különbségek pedig arra utalnak, hogy különböző beszélők esetén más ismertetőjegyek lehetnek fontosak. A 3. ábra alapján NG-t a glottalizáció, míg FG-t az alapfrekvencia-menet alapján volt könnyebb azonosítani.

A hallgatók az esetek 76%-ában ugyanarra a kérdésre fordított lejtárszási sorrendben legfeljebb egy skálaértékkel eltérő választ adtak. Ugyanez a megbízhatósági mérték az első pártípusra 83% volt. Ez azt jelenti, hogy a kísérleti személyek következetesen válaszoltak a kérdésekre, és ha újra elvégeznénk a kísérletet, valószínűleg az egyes pártípusokra hasonló eredményeket kapnánk.



3. ábra

Találati arány az egyes bemondókra és pártípusokra

Összefoglalás és kitekintés

Percepció kísérlettel vizsgáltuk meg, hogy vajon a végső glottalizáció érzékelhető-e a hallgatók számára, és hogy szerepet játszik-e a beszélő személy felismerésében. Mivel a beszélőfelismerési folyamatban felhasznált ismertetőjegyek beszélőnként és hallgatónként eltérőek (van Lancker et al. 1985), nem volt célunk e szerep általános érvényű bizonyítása: hipotézisünk szerint ez csak egyes beszélőkre és egyes hallgatókra korlátozódik. Az eredmények alapján ezekkel a megkötésekkel a fenti kérdésekre igen a válasz, de a következtetések hitelességéhez további kísérletek elvégzése szükséges. Az eredmények biztatóak, mert még egy olyan erőteljes beavatkozás mellett, mint az alaphérfekvencia-görbe átültetése, is kimutatható volt a végső glottalizáció hatása.

Feltételezéseinknek megfelelően a beszélőfelismerési folyamatban az alaphérfekvencia-menetnél általában kisebb súlyú, de szignifikáns ismertetőjegy a glottalizáció. A négy beszélő közül a gyakran glottalizáló nő esetén azonban a végső glottalizáció több információt hordozott a beszélő személyéről, mint az alaphérfekvencia-menet.

Az eredményeket szintetikus beszédre kaptuk, így bizonyos szintű fenntartásokkal kell kezelnünk azokat. Bár a kísérlet második és harmadik része alapján ezek minősége (a beszélőfelismerés szempontjából) megközelítette a természetes beszédet, nem biztosított, hogy ez a felvételek minden fontos jellegzetességét utánozni tudta.

Az eredmények egyrésztől hozzájárulnak a beszéd paralingvisztikai tulajdonságainak megértéséhez, másrésztől alkalmazhatók a beszédtechnológiában.

Irodalom

- Allen, J. Sean – Miller, Joanne L. 2004. Listener sensitivity to individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America* 115. 3171–3183.
- Gordos Géza – Takács György 1983. *Digitális beszédfeldolgozás*. Műszaki Könyvkiadó, Budapest.
- Gósy Mária 2004. *Fonetika, a beszéd tudománya*. Osiris Kiadó, Budapest.
- Grassegger, Hans 2003. A hangminőség szubjektív és objektív megítélése. *Beszéd kutatás* 2003. 44–56.
- Henton, Caroline G. – Bladon, Anthony 1987. Creak as a sociophonetic marker. In Hyman, Larry M. – Li, Charles N. (eds.): *Language, speech and mind: Studies in honour of Victoria A. Fromkin*. Routledge, London, 3–29.
- Huber, Dieter 1992. Perception of aperiodic speech signals. *Proceedings of International Conference on Spoken Language Processing (ICSLP) '92*. Alberta, Canada, 503–506.
- Klatt, Dennis H. – Klatt, Laura C. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* 87. 820–857.
- Kreiman, Jody – Gerratt, Bruce R. – Precoda, Kristin – Berke, Gerald S. 1992. Individual differences in voice quality perception. *Journal of Speech, Language, and Hearing Research* 35. 512–520.
- Laver, John 1980. *The phonetic description of voice quality*. Cambridge University Press, Cambridge.
- Markó Alexandra 2005. *A spontán beszéd néhány szupraszegmentális jellegzetessége*. PhD-értekezés. ELTE, Budapest.
- Pardo, Jennifer S. – Remez, Robert E. 2005. *The perception of speech*. Technical Report. Speech Perception Laboratory, Barnard College, New York.
- Redi, Laura – Shattuck-Hufnagel, Stefanie 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29. 407–429.
- Slifka, Janet 2000. *Respiratory constraints on speech production at prosodic boundaries*. PhD-disszertáció. MIT, Massachusetts.
- Van Lancker, Diana – Kreiman, Jody – Wickens, Thomas D. 1985. Familiar voice recognition: patterns and parameters. Part II. *Journal of Phonetics* 13. 39–52.

A szerző ezúton fejezi ki köszönetét Kenneth N. Stevens és Stefanie Shattuck-Hufnagel (MIT, USA) szakmai útmutatásáért és Maricza Istvánnak az adatok elemzéséhez nyújtott segítségért. A munka anyagi háttérét a Fulbright-ösztöndíj biztosította.

BESZÉDADATBÁZIS PROZÓDIÁJÁNAK SZEREPE A GÉPI BESZÉD HANGZÁSÁBAN ÉS ÉRZELMI TARTALMAK KIFEJEZÉSÉBEN

Vidám, avagy szomorú a beszédszintetizátor?

Zainkó Csaba – Fék Márk

Bevezetés

A beszédelemek összefűzésén alapuló szövegfelolvasó rendszerek természetes emberi bemondásokból kivágott hullámformarészleteket használnak a gépi beszéd előállítására. Ezeket többnyire egy hullámforma-elemtárban tárolják. A gépi beszéd előállításakor az itt tárolt hullámformákat fűzzük össze, majd az így előállított nyers beszéden jelfeldolgozási módszerek segítségével megvalósítjuk a kívánt prozódiaát. Prozódia alatt az alapfrekvencia-, hangidőtartam- és intenzitásértékek sorozatát értjük, amely alapján a szintetizátor ki tudja alakítani a hangsúlyozást, a beszéddallamot és a beszédritmust. A hullámforma-adatbázis készítéséhez szólistát olvas fel egy bemondó érzelmi töltés nélkül, állandó alapfrekvenciával. Ez a felolvasási módszer azzal indokolható, hogy az elemek tisztább artikulációja növeli az összefűzött beszéd érthetőségét, az azonos hangmagasság segíti a jelfeldolgozó algoritmusok működését, illetve valamelyest folytonossá teszi az összefűzött beszéd hangzását.

Célkitűzés

A jelen tanulmányunkban azt kutatjuk, hogy az aktuálisan használt beszédfeldolgozási és szintézistechnikáink milyen mértékben alkalmasak érzelmi töltetű gépi beszéd előállítására. Ezek mellett azt is megnézzük, hogy a korábbi évek során általánosan használt beszédszintézis-rendszerünknek a semlegetstől eltér-e az érzelmi megítélése.

A fenti kérdések megválaszolásához azt elemezzük, hogy a szintetizált beszédben megjelenik-e az érzelem, ha érzelmi töltéssel mondatjuk be a hullámforma-adatbázis készítéséhez szükséges szólistát. Továbbá azt is megvizsgáljuk, hogy a hullámforma-adatbázis érzelmi töltete mellett, a prozódiaának mekkora szerepe van a szintetizált mondat érzelmi megítélésében.

Korábbi kutatások

Az ember által kifejezhető érzelmek és érzelmi árnyalatok száma több száz (Cowie 2000). Ezek közül a témával foglalkozó pszichológiai szakirodalom megkülönböztetett figyelmet szentel a következő hat, ún. alapérzelemnek: öröm,

harag, szomorúság, félelem, meglepődés és undor. A beszédbeli érzelmek kifejezést vizsgáló kutatások is legtöbbször ezen alapérzelmek vizsgálatára összpontosítanak. Az érzelmi töltés a beszéd akusztikai szintjén is megjelenik, tehát a beszélő hanglejtéséből és hangszínezetéből következtethetünk érzelmi állapotára.

A természetes és gépi beszéd érzelmek kifejező képességét vizsgáló kísérletek kimutatták, hogy a beszédben megjelenő érzelmi töltést nem kizárólagosan a hangmagasság, az artikulációs sebesség és a hangintenzitás hordozza, hanem abban a beszéd hangszíneképe is jelentős szerepet játszik (Schröder 1999). Ezt a magyar nyelv esetében is alátámasztotta egy korábbi vizsgálatunk (Fék et al. 2005). Ennek során semleges szemantikai tartalmú, színészek által eljátszott 9 különböző érzelmet hordozó mondatot vizsgáltunk. A prozódiaátültetési kísérletünk során az érzelmi töltetet hordozó mondatok alaphangfrekvencia-menetét és hangidőtartamait semleges mondatokra másoltuk. A felismerési tesztek alapján a két tényező leginkább a meglepődés, az értetlenség és a lekicsinylés kifejezésében játszik szerepet. Egy későbbi kísérletben az előbbi két paraméter mellett az intenzitásmenetet (Szabó 2005) is átültettük a semleges mondatokra, ami a harag és a szomorúság felismerhetőségét növelte. Ugyanakkor a félelem és az öröm kevésbé volt felismerhető, ezért feltehető, hogy ezen érzelmek kifejezésében a hangszíneképnek fontos a szerepe.

Az elem-összefüzéses beszéd szintetizátor által előállított gépi beszéd hangszíneképét utólagosan jelfeldolgozási eljárásokkal módosíthatjuk. Jelenleg ezek a módszerek egyrészt nem elég rugalmasak, másrészt az így módosított beszéd hangzása sem kielégítő. A másik lehetőség, hogy már a beszédatadatbázis készítésekor többféle beszédet veszünk fel, és az összefüzés során mindig a megfelelő hangszíneképű elemek közül választunk. Ilyen vizsgálatot végzett Schröder német nyelvre (2003).

Bulut és munkatársai (2002) szintén vizsgálták az érzelmek kifejezését a szintetizált mondatokban. Első lépésben 5 semleges tartalmú mondatot kerestek, majd meghatározták az ezek előállításához szükséges diád elemeket (a diád két félhangból álló elem). Ezek után 4 érzelmi változatban (semleges, öröm, bánat, harag) egy-egy szöveget készítettek. A szövegek mindegyike tartalmazta az érzelmi előállításához szükséges diád elemeket, emellett az egyes szövegek mindig az adott érzelmet felidéző mondatokból álltak. A szöveget egy félprofi színész olvasta fel. Egy elem-összefüzésen alapuló beszéd szintetizátor segítségével előállították az 5 semleges mondatot mind a 4 érzelmi töltetű adatbázissal, ami 20 mondatot jelentett. Egy szintetizált mondat alaphangfrekvencia-menetét, hangidőtartamait és intenzitásmenetét mindig a megfelelő érzelmű töltettel felvett eredeti mondatról ültették át. A 20 mondathoz további 60 mondatot is készítettek oly módon, hogy a fenti prozódiai paramétereket más-más érzelmi változatú mondatokról ültették át. Így például örömteli adatbázisból, de haragos prozódiamenettel is készült mondat. A meghallgatási tesztek folyamán mind a 4 érzelmet sikerült a véletlen találgatást jóval meghaladó mértékben azonosítani egy 4 választási lehetőséget adó tesztben. Ugyanakkor az öröm felismerési szá-

zaléka körülbelül csak a fele volt a másik három változatra kapott eredménynek. A különböző adatbázisok és prozódiamenetek keverésével előállított mondatok mindig kevésbé voltak felismerhetők, mint az egymásnak megfelelő párosítás esetén.

Anyag és módszer

A vizsgálathoz kétfajta beszéd szintetizátort használtunk, hullámforma-összefűzéses és szinuszosmodell-alapú szintetizátort. A hullámforma-összefűzéses beszéd szintetizátor hagyományos fajtája, mint a magyar nyelvű Profivox (Olaszy et al. 2000), a beszédet ún. diád és triád hullámformákból állítja elő. A diád hullámforma két félhangot tartalmaz, a triádok egy fél, a hozzá csatlakozó egész és a következő fél hangból állnak. A beszéd előállítása során a hangsorozatnak megfelelően kiválasztott elemeket fűzi egymás után a rendszer. Az így kapott nyers beszédjelre a rendszer ráülteti a szabályok alapján az eredeti szövegből előállított prozódiaiát. Ennek során a hangok időtartama, alaphangfrekvenciája és energiája módosul.

A szinuszosmodell-alapú beszéd szintetizátor (Stylianou et al. 1997) ugyanaból az adatbázisból indul ki, mint a hullámforma-összefűzéses szintetizátor, de a jelet nem időtartományban módosítja, hanem felharmonikusokra bontja, majd a szükséges módosításokkal generált felharmonikusokat újból összegezve állítja elő a gépi beszédet. A modellben a nehézséget a felharmonikusok fázisának kezelése jelenti. Ezekből a hibákból adódik a szintetizátor jellegzetes fémes hangzása. A modell nagy előnye, hogy az alaphangfrekvenciát és a hangok időtartamát folytonosan és pontosan be lehet állítani. A hullámforma-összefűzéses szintetizátor esetében erre csak korlátozottan van lehetőség (főleg az alaphangfrekvencia esetében).

A diádokat és a triádokat természetes emberi bemondásokból vágjuk ki, és tároljuk el a hangadatbázisban. A diád és triád elemeket tartalmazó hangadatbázis készítése során egy bemondó egy olyan listát olvas föl, amelynek minden egyes szava vagy logatomja (értelem nélküli hangsorozat) egy adott diádot vagy triádot tartalmaz.

A hangadatbázisok felvétele

A kísérletünk a következő mintamondaton alapult: „A menüben minden szükséges információ elhangzik.” A mondat szintetizálásához szükséges diád és triád elemek elkészítéséhez szavakat és logatomokat mondott fel egy színész (44 éves) és 2 színésznő (29, 30 évesek) érzelmi töltet nélkül. A semleges forma mellett a következő 6 érzelmi töltettel is felolvasták a színészek a listát: kedvesség, szomorúság, félelem, undor, harag és meglepettség, amelyek a kedvesség kivételével az alapérzelmeknek felelnek meg. A színészek az érzelmeket az instrukciók szerint eljátszották, a felvételek tehát nem átélt érzelmeket továbbítanak.

A feldolgozás további menete is azonos volt azzal az eljárással, amelyet a korábbi évek során a Profivox szintetizátor hangadatbázisának készítésekor használtunk.

A bemondók a mintamondatot is felolvasták a szavakhoz és a logatomokhoz hasonlóan a 6-féle érzelmi töltettel és semleges formában is.

A tesztmondatok előállítása

A tesztmondatok előállítása eltérő módon történt a kétféle szintetizálási eljárás esetében. A hullámforma-összefüzeses szintetizátor rendelkezik saját prozódia generáló modullal, amely a bemeneti szövegből szabályalapon előállítja az idő-, alapfrekvencia- és energiaadatokat. Mivel a prozódia előállításához a bemeneti szövegből indul ki, ezért minden adatbázishoz egységesen ugyanazt a prozodiát generálja. Változatosság abban jelenik meg, hogy az előírt alapfrekvencia-menetet nem konkrét frekvenciaértékként adja meg, hanem csak azt, hogy egy alapértékhez képest mennyivel módosítandó a nyers beszédjel. A rendszer alapértékként a felvett hangadatbázis átlagos alapfrekvenciáját használja. A szintetizátor lehetőséget ad egyes prozódiai paraméterek bizonyos mértékű módosítására. Változtatni tudjuk az alaphangmagasságot és az átlagos beszédtempót az adatbázis átlagos alaphangmagasságához és beszédtempójához képest. Mivel a hangadatbázisok különböztek a fenti két értékben, ezért ezeket a beállításokkal megpróbáltuk kiegyenlíteni, hogy az összehasonlítások során ezek az eltérések ne befolyásolják a hallgatót.

A szinuszos szintetizátor nem rendelkezik prozódiai egységgel, ezért két külső forrás adatait használta fel. Az egyik a hullámforma-összefüzeses szintetizátor által előállított, a másik pedig természetes emberi bemondásokból származtatott prozódia. Az emberi beszédből származtatott prozódia esetén a természetes mondat legtöbb paraméterét próbáljuk változatlan formában megőrizni a szintetizálás során. Erre az üzemmódra a hullámforma-összefüzeses szintetizátor esetében sajnos jelenleg nincs lehetőség, ezért az ilyen jellegű tesztek csak a szinuszos szintetizátorral tudtuk elvégezni.

A meghallgatásos tesztek alátámasztására statisztikai analízist végeztünk. Ennek során azt vizsgáltuk, hogy az egyes tesztekben a tesztelők szignifikánsan a véletlen tippelés szintje fölélt jelöltek-e meg egy adott kategóriát. A vizsgálatokhoz minden esetben a binomiális próbát alkalmaztuk 95%-os konfidenciaszint mellett. A táblázatokban *-gal jelöltük a sikeres tesztek, azaz amikor $p < 0,05$. Az adatok értékelése a Matlab 6.0 programmal történt.

A kísérletek

A teszt több részből állt, átlagosan 17 perc alatt végeztek a tesztelők a teljes kísérlettel. A tesztet weben keresztül lehetett meghallgatni, és a válaszokat is ugyanott kellett megadni. A hangbejátszásokat tetszőleges alkalommal hallgathatták meg a tesztelők, de a válaszok után visszalépésre nem volt lehetőség. A tesztet 23 fő hallgatta meg, 12 férfi és 11 nő (20 és 60 év között). A teszt egyes részein belül a hanganyagok lejátszása véletlen sorrendben történt.

A bemondások ellenőrzése

A teszt ezen részében a bemondott szavakat és logatomokat kellett a hallgatóknak értékelniük az érzelmi tartalom szerint. Ez azért fontos, hogy a belőlük

készített hangadatbázisok érzelmi minőségét meg tudjuk határozni. A 7 lehetséges választási lehetőség a következő volt: semlegesség, kedvesség, harag, szomorúság, meglepődés, undor, félelem.

Minden érzelemből és a semleges bemondásból is egy-egy szót és két-két logatomot választottunk ki a következők szerint: az első logatom olyan, amelyből hangsorkezdő és hangsorbefejező elemet használt fel a szintetizátor: *mám*. A második logatom egy magánhangzót is tartalmazó diád vagy triád elemhez szükséges bemondás volt: *amena*. A harmadik pedig egy mássalhangzó-kapcsolódást tartalmazó szó: *peremfeltétel* (vagy *peremfelvétel*). Ebben az esetben azt is vizsgáltuk, hogy a nagyobb odafigyelést kívánó szó kimondása során mennyire tudta a bemondó a megfelelő érzelmet is kifejezni. A stúdióban a felolvasási listán minden esetben a *peremfeltétel* szó szerepelt, de a felolvasó színész sok esetben tévesztett, és *peremfelvételt* olvasott. Mivel a kívánt mássalhangzó-kapcsolódás a szó közepén nem változott, ezért az ismételt felvételre nem volt szükség, ugyanakkor kitűnt, hogy ez a szó nehezebben olvasható fel a többihez képest. Az első részben így 21 meghallgató szó és logatom szerepelt.

A bemondó a harag és a meglepődés érzelmet játszotta el a leghitelesebben, 90% és 84% ítélte az elvárt érzelmi töltésűnek a szavakat (1. táblázat). A félelem töltetű szavak esetében 58%-os átlag jött ki annak ellenére, hogy az *amena* logatomot majdnem 65% kedvesnek ítélte, és csak 22% azonosította félelemként. A legrosszabbul bemondott szó az undorral bemondott *peremfelvétel* lett, amelynél az undorra csak egyetlen tesztlő asszociált. A többi undort kifejező logatom 50% felett teljesített. A szomorú és a semleges szavakat és logatomokat nehezen azonosították a hallgatók, többnyire egymással cserélték fel a két érzelmet. A kedves szavak és logatomok bemondása mutatta a legnagyobb szórást, legtöbbször a semlegesre és a meglepettre gondoltak, de az egyik vizsgált szót 20%-ban haragosnak vélték.

1. táblázat: A szavak és logatomok érzelmi töltöttségének tesztlők általi azonosítása (* = az érték szignifikáns)

| | <i>amena</i> | <i>mám</i> | <i>peremfelvétel</i> <i>peremfeltétel</i> | Átlag |
|--------------------|--------------|------------|--|-------|
| Félelem | 22% | 74%* | 78%* | 58%* |
| Harag | 100%* | 83%* | 87%* | 90%* |
| Kedvesség | 39%* | 17% | 39%* | 32%* |
| Meglepődés | 78%* | 100%* | 74%* | 84%* |
| Szomorúság | 26% | 57%* | 61%* | 48%* |
| Semlegesség | 74%* | 30%* | 26% | 43%* |
| Undor | 52%* | 61%* | 4% | 39%* |
| Átlag | 56%* | 60%* | 53%* | 56%* |

Kimutatható, hogy a *mám* logatom esetében általában nem asszociáltak a hallgatók a *lám* meglepődést kifejező szóra. Az adatok alapján az összes megle-

pódéstől eltérő érzelmű *mám* logatom esetén, az eseteknek csak 4,3%-ában jelölték meg a hallgatók a meglepődést. A meglepődéses érzelmű *mám* logatom esetén viszont lehetséges, hogy ez az asszociáció segítette a 100%-os azonosítást.

Az adatok alapján nem állapítható meg, hogy bármelyik szó vagy logatom segítségével könnyebb vagy nehezebb lenne egy adott érzelmet eljátszani. Az eredmények ismeretében, a felvételek újbóli meghallgatása után úgy találtuk, hogy a bemondóknak nem az adott érzelem eljátszása okozta a nehézséget, hanem az, hogy az adott érzelmet folyamatosan, minden szóra és logatomra megvalósítsák. A szavakat és logatomokat átlagosan 56%-ban azonosították megfelelő érzelmű töltésűnek a hallgatók. Ez az érték alacsony, de a kísérlet szempontjából a felvétel körülményei külső adottságok, ezért ezen nem tudtunk változtatni. A szintetizált mondatok értékelésekor figyelembe vettük, hogy az adott mondathoz felhasznált hangadatbázis szavai és logatomjai érzelmi szempontból mennyire voltak jól azonosíthatók.

A teszt másik részében a színész teljes mondatos bemondásait ellenőriztettük a hallgatókkal. Itt szintén a szavak és logatomok esetében is használt 7 lehetőség közül lehetett választani. A teljes mondatok közül a félelem, a harag, a meglepődés, a szomorú és a semleges bemondás gyors változatát egyértelműen azonosították a hallgatók (2. táblázat).

2. táblázat: A mondatok érzelmi töltöttségének tesztelők általi azonosítása
(Az összes érték szignifikáns.)

| Semleges lassú | Semleges gyors | Kedves | Harag | Szomorú | Meglepődés | Undor | Félelem |
|-------------------|-------------------|--------|-------|---------|------------|-------|---------|
| 48% | 87% | 70% | 96% | 83% | 96% | 78% | 87% |

A kedves bemondás esetén a tesztelők válaszaiban 30% volt a semleges jelölés. A lassú semleges bemondásnál 25-25% jelölte meg a kedves és a szomorú érzelmet a semleges mellett. A tesztalanyok nem ismerték a bemondót, és a teszt ráhangolási szakaszt sem tartalmazott, ahol hozzászokhattak volna a beszélőhöz, ezért a semleges érzelmi állapotát sem tudták egységesen megítélni. Az undor érzelem azonosítása – a szavakhoz és a logatomokhoz hasonlóan – a mondatoknál is alacsony volt. A felvételek készítésekor a színész is jelezte, hogy ennek eljátszása nehéz számára, tehát valószínűleg a bemondások hibája okozta az alacsony azonosítási arányt.

A szintetizált mondatok értékelése

A harmadik részben a különböző érzelmi töltetű és a semleges adatbázisokból előállított mondatokat vizsgáltattuk, ahol a választási lehetőségek azonosak voltak az előző részekkel. A mondatok prozódijája minden esetben a hullámformaösszefűzéses szintetizátoré, a szintézistechnológiát tekintve 16 mondatból 11-et a Profivox, 5-öt a szinuszos modell generált. Mivel a prozódia szabályalapon generált, ezért itt az adatbázisok hangszínezetének hatását vizsgáltuk.

A 3. táblázatban azoknak a mondatoknak az értékelését találjuk, amelyek a semleges hangadatbázis felhasználásával, de különböző alapfrekvencia-beállítással készültek. A természetes módon bemondott hangadatbázisból szintetizált mondatot ugyanannyian ítélték semleges érzelműnek a normál alapfrekvenciájú esetekben, mint a monoton bemondásokból készült hangadatbázis esetén. Ez azt jelenti, hogy a mondat semleges érzelmi megítélése szempontjából nincs jelentősége annak, hogy a bemondó azonos alapfrekvencián vagy a természetes dallammenettel mondta be a szavakat és a logatomokat. A Profivox szintetizátor alapváltozata monoton dallammenetű adatbázissal működik, amely esetében az előállított mondatot a semleges mellett szomorúnak (43%) ítélték a tesztelők.

3. táblázat: Különböző alapfrekvenciájú, természetes hangadatbázissal szintetizált mondatok érzelmi megítélése (* = az érték szignifikáns)

| Hangmagasság | Szintetizátor | Leggyakrabban felismert érzelmek |
|------------------------|---------------------------|--|
| alacsony F_0 | összefüzéses szinuszos | szomorú 48%*, semleges 30%* semleges 48%*, szomorú 26%* |
| normál alatti F_0 | összefüzéses | semleges 78%*, szomorú 13% |
| normál F_0 | összefüzéses szinuszos | semleges 52%*, kedves 17%, szomorú 17% semleges 74%* |
| magas F_0 | összefüzéses | félelem 26%, kedves 26%, semleges 26% |
| normál F_0 (monoton) | összefüzéses | semleges 52%*, szomorú 43%* |

A 4. táblázat az érzelmi töltetű mondatok esetében leggyakrabban felismert érzelmeket mutatja. A hangadatbázis készítéséhez használt szavak és logatomok érzelmi töltete befolyással van a kész mondatokra, ezért a jobb összehasonlíthatóság érdekében a táblázat utolsó oszlopában feltüntetjük a szavak és logatomok érzelmi töltetének átlagos értékeit is.

4. táblázat: Különböző érzelmi hangadatbázissal szintetizált mondatok érzelmi megítélése (* = az érték szignifikáns)

| Érzelmi töltés | Szintetizátor | Leggyakrabban felismert érzelmek | Szavak és logatomok |
|----------------|------------------------|--|---------------------|
| harag | összefüz. szinuszos | harag 65%*, semleges 22% semleges 61%*, harag 35%* | 90%* |
| félelem | összefüz. | félelem 39%*, szomorú 29% | 58%* |
| kedvesség | összefüz. szinuszos | kedves 39%*, semleges 26%, harag 26% semleges 61%*, kedves 17% | 32%* |
| szomorúság | összefüz. szinuszos | szomorú 52%*, kedves 26%, semleges 22% szomorú 57%*, semleges 22%, kedves 17% | 48%* |
| meglepődés | összefüz. | félelem 61%*, meglepődés 22% | 84%* |
| undor | összefüz. | harag 57%*, semleges 17%, undor 13% | 39%* |

A 3. és a 4. táblázatokból látható, hogy a harag nem jelenik meg a semleges adatbázisokkal készült mondatokban, viszont a harag (65%) és undor adatbázisokkal (57%) készültekben igen. A szomorúság leginkább a szomorú adatbázis esetén észlelhető (52%), de a semleges változatokban, különösen az alacsony F_0 -al előállított esetben (48%) is jelen van. Azonos prozódia mellett a hangadatbázis érzelmi töltete a harag és a szomorúság esetében jön át egyértelműen.

A félelem leginkább a meglepődés adatbázissal észlelhető (61%), míg a félelmet tartalmazó adatbázissal szintetizált mondatban kevésbé van jelen (39%), amit a kiinduló hangadatbázis hibája is okozhat. Ugyanakkor a magas alaphéremvenciájú semleges mondat is valamennyire kifejezi a félelmet (26%). A kedvesség leginkább a kedves adatbázissal készült mondatban észlelhető (39%), ugyanakkor két semleges változatban (17%, 26%) és a szomorú adatbázis (22%) esetén is megjelenik. Tehát a félelem és a kedvesség esetében a hangadatbázis érzelmi töltete megjelenik a szintetizált mondatokban, de az egyértelmű azonosításhoz nem elegendő.

A meglepődést csak a hozzá tartozó adatbázis esetén észlelték (22%), de ott is jóval megelőzte a félelem (61%). A meglepődést csak kismértékben azonosították, annak ellenére, hogy a szavakat és logatomokat 84%-ban megfelelően felismertek a tesztelők. Az undor kizárólag az undor adatbázis esetén jelenik meg, és ott is csak kismértékben (13%), mivel az undor adatbázissal készült mondatot elsősorban haragnak azonosították (57%). A meglepődés és az undor esetében a hangadatbázis érzelmi töltete nem jelenik meg szignifikánsan a mondatokban.

A prozódia szerepe az érzelmi töltetű mondatokban

Az utolsó részben megvizsgáltuk, hogy a prozódiaának vagy a hangadatbázis érzelmi töltetének van-e nagyobb szerepe a mondatok érzelmi töltésének megítélésénél. Továbbá azt is vizsgáltuk, hogy ambivalens párosítások esetén (például: szomorú adatbázis – haragos prozódia) milyen érzelmet hall ki a tesztelő a mondatból. A tesztben csak szinuszos modellel generált mondatok szerepeltek. A választási lehetőség az eddig is használt 7 érzelem volt. A generált mondatok adatbázisai a következők: semleges, kedves, szomorú és haragos. A prozódia természetes emberi bemondásokból származott. A prozódia az adatbázisoknál használt négy érzelmenek felelt meg, így az összes lehetséges variációt előállítva 16 mondatot kaptunk.

A természetes prozódiajú szintetizált mondatok között a harag érzelem volt a legdominánsabb. Ha az adatbázis haragos volt, akkor a prozódia beállításától függetlenül a haragos érzelmet jelölték meg a legtöbben (5. táblázat). Haragos prozódia és más adatbázis esetén is legtöbbször a haragosra asszociáltak a tesztelők, kivéve a szomorú hangadatbázis esetében. Ebben az esetben a hangadatbázis és a prozódia érzelmi töltésétől is eltérő új érzelemet, a félelmet jelölte meg a hallgatók 57%-a.

Abban az esetben, amikor a párosításokban azonos érzelmek szerepeltek prozódiaiként és hangadatbázisként, a tesztelők mindig a megfelelő érzelmet választották – hasonlóan Bulut kísérleteihez (Bulut et al. 2002). A kedves adatbázis

esetén a szomorú prozódia nem érvényesült, 48%-ban semleges érzelmet választottak a tesztelők. A hangadatbázisok közül a szomorú csak a szomorú érzelm esetén tudta megőrizni érzelmi töltöttségét, minden más prozódia esetén eltérő érzelmet jelölt meg a hallgatók többsége. Ennek oka a kísérletből egyértelműen nem határozható meg, mert a szomorú adatbázis kiinduló adatát jelentő szavakat és logatomokat is csak 48%-ban azonosították megfelelően a hallgatók.

5. táblázat: A leggyakrabban kiválasztott érzelmek a prozódia és hangadatbázis függvényében (* = az érték szignifikáns)

| Prozódia | Harag | Kedves | Szomorú | Semleges |
|---------------------------|---------------------------|--|---|------------------------------|
| Haragos adatbázis | harag 91%* | harag 48%* semleges 30%* | harag 39%* semleges 39%* | harag 74%* semleges 22% |
| Kedves adatbázis | harag 70%* undor 17% | kedves 57%* semleges 30%* | semleges 48%* kedves 35%* | semleges 44%* kedves 35%* |
| Szomorú adatbázis | félelem 57%* harag 26% | semleges 48%* kedves 35%* | szomorú 65%* semleges 22% | semleges 44%* kedves 35%* |
| Semleges adatbázis | harag 57%* undor 30%* | semleges 74%* kedves 13% undor 13% | semleges 57%* szomorú 22% undor 22% | semleges 74%* |

Következtetések

Az eddig használt hangadatbázis-rögzítési eljárások és hullámforma-összefüzeses szintetizálási technikáink csak korlátozottan alkalmasak érzelmi töltetű mondatok előállítására. Az érzelmű töltetű felvételek készítése során követett eljárás nem biztosítja a megfelelő minőségű érzelmi hangadatbázis létrehozását. A hullámforma-összefüzeses szintetizátor parametrikus prozódia generáló eljárása szintén csak korlátozottan alkalmas érzelmi szintetizálásra.

A korábbi évek során használt Profivox szintetizátor alapváltozata részben semleges, részben szomorú érzetet kelt a hallgatókban.

A hangadatbázis érzelmi töltete harag és szomorúság esetében egyértelműen, míg meglepődés esetében egyáltalán nem jelent meg a szintetizált mondatban. Kedvesség, félelem és undor esetében a kiinduló hangfelvételek hibája miatt egyértelmű következtetéseket nem tudunk levonni.

A szintetizált mondat érzelmének megítélésében az érzelmi töltetű mondatokból származtatott prozódia hatása függ a hangadatbázis érzelmi töltésétől. A haragos prozódia meghatározó volt a vizsgált érzelmek közül. Ambivalens prozódia-hangadatbázis párosítás esetén új érzelem is kialakulhat.

Irodalom

- Bulut, Murtaza – Narayanan, Shrikanth S. – Syrdal, Ann K. 2002. Expressive speech synthesis using a concatenative synthesizer. In: *Proceedings of the 7th International Conference on Spoken Language Processing*. Denver, 1265–1268.
- Cowie, Roderick 2000. Describing the emotional states expressed in speech. In: *Proceedings of International Speech Communication Association Tutorial and Research Workshop on Speech and Emotion*. Newcastle, 11–18.
- Fék Márk – Olasz Gábor – Németh Géza – Gordos Géza 2005. Érzelem kifejezése gépi beszéddel. *Beszédkutatás* 2005. 134–144.
- Olasz Gábor – Kiss Géza – Németh Géza – Olasz Péter 2000. Profivox: a legkorszerűbb hazai beszéd szintetizátor. *Beszédkutatás* 2000. 167–179.
- Schröder, Marc 1999. Can emotions be synthesized without controlling voice quality? In: *Phonus 4*. Research Report of the Institute of Phonetics, Saarland University. Saarbrücken, 37–55.
- Schröder, Marc – Grice, Martine 2003. Expressing vocal effort in concatenative synthesis. In Solé, Maria-Josep – Recasens, Daniel – Romero, Joachim (eds.): *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona 3-9 August 2003*. Universitat Autònoma de Barcelona, Barcelona, 2589–2592.
- Stylianou, Yannis – Dutoit, Thierry – Schroeter, Jürgen 1997. Diphone concatenation using a harmonic plus noise model of speech. In: *Proceedings of Eurospeech '97*. Rhodes, 613–616.
- Szabó János 2005. *Érzelmi beszédatadátbázis feldolgozása és összehasonlító elemzése*. Diplomaterv. BME, Budapest.

A kutatást az NKFP 2. programja (szerződés szám: 2/034/2004) támogatta.

ÚJ ELJÁRÁS A GÉPI BESZÉDFELISMERÉS KÖRNYEZETFÜGGŐ BESZÉDHANGMODELLJEINEK KIALAKÍTÁSÁRA

Mihajlik Péter – Fegyő Tibor – Tatai Péter

Bevezetés

A gépi beszédfelismerés hatásfokát alapvetően befolyásolja az alkalmazott beszédhangmodellek kialakításának módja. A legegyszerűbbek az ún. monofón modellek, amelyeknél fonémaként egyetlen beszédhangmodellt használunk. E megközelítés előnye nyilvánvaló, hiszen általában alig negyven beszédhangmodell elegendő az európai nyelvek beszédakusztikájának teljes lefedésére. A hátránya is e helyen keresendő: igen nehéz a beszéd óriási változatosságát (Olaszy 2003) ilyen kisszámú beszédhangmodellel leírni.

Bár az egyes beszédhangmodellek komplexitásának növelése segíthet a beszédfelismerési pontosság növelésében, az általános tapasztalat mégis az, hogy nagyobb számú modell alkalmazása teszi alapvetően használhatóbbá, pontosabbá a beszédfelismerő rendszereket (Young 2000). A legelterjedtebb megközelítés a beszédhangok fonetikai-fonológiai környezetének figyelembevételével történő modellalkotás. Ezen belül szinte egyeduralkodó az a szemlélet, amely szerint csak az adott beszédhangot közvetlenül megelőző és közvetlenül követő beszédhangot mint fonetikai környezetet veszik figyelembe. Ezt ún. trifón modellezésnek szokták hívni. Például a *pamacs* szó első [ɔ] hangját a $p-a+m$ (általánosított) trifón modellezi, ahol a p - a bal kontextust, a $+m$ a jobb kontextust jelöli. Fontos, hogy maguk a kontextusok fizikailag nem tartoznak a trifónhoz, csak a modellezett hang.

A trifón modellezés előnye, hogy mind a fonetikai, mind a fonológiai koartikulációs jelenségek többségét – definíciójából eredően – jól képes kezelni (Jurafsky 2001), és ez a felismerési pontosságban is tükröződik. Hátránya viszont – szintén a definíciójából adódóan –, hogy jóval bonyolultabb módon lehet általános felismerési hálózatokat építeni belőlük, mint monofón modellekből. További nehézséget okoz a modellek finomszerkezetének, paramétereinek meghatározása, vagyis az akusztikus modell tanítása.

A továbbiakban a trifón modellezés problémáival, szokásos és újszerű megoldásaival, valamint az ezekkel elért beszédfelismerési eredmények összehasonlító elemzésével foglalkozunk.

A trifón modellezés problémái

a) Tanítás

Problémákat vet fel, ha a környezetfüggő beszédhangmodelleknél minden fonológiai környezetet megkülönböztetünk, azaz **általánosított trifón modelleket** használunk. Ez könnyen szemléltethető a következő példával.

Ha egy adott nyelv fonémáinak számát $P = 40$ -nek tekintjük, akkor a fonológiai szabályszerűségeket figyelmen kívül hagyva, elméletileg P^3 , azaz 64 000 környezetfüggő beszédhangmodellt kapunk, hiszen a 40 beszédhang mindegyike elvileg 40×40 -féle fonológiai környezetben valósulhat meg. Kimutatható, hogy még a hangsorépítési szabályok (Siptár 1995) figyelembevételével is több ezer különböző általánosított trifón modell adódik. Elméletileg ugyan a modellek nagy száma kedvező a beszédváltozatosság lefedése szempontjából, a gyakorlatban azonban több problémát vet fel.

Ugyanis minden beszédhangmodell betanításához legalább néhány száz mintára van szükség a modellezett hangból. Ez a fenti példában szereplő általánosított trifón modellek esetén azt jelenti, hogy még a legritkább általánosított trifón modellnél, mint pl. az $a-g+ny$ esetén is több száz, lehetőleg természetes beszédben előforduló példát kellene használnunk a tanítási eljárás során. Sőt ha beszélőfüggetlen felismerőt szeretnénk, arra sem árt ügyelni, hogy a beszélők nemben, korban változatosak legyenek.

Látható, hogy amennyiben minden fonológiai környezetet különbözőnek tekintünk, a trifón modellek betanításához olyan méretű és minőségű beszédadatbázisra lenne szükségünk, amely a világon szinte sehol nem áll rendelkezésre.

A megoldást az egyes (általánosított) trifón modellek fonetikus környezetük alapján történő összevonása, illetve a modellek számának az adott beszédadatbázishoz illesztése jelenti. Ilyen megoldás például az általunk kidolgozott fonológiai csoportosítás módszere, valamint a széles körben használt fonetikus döntési fákon alapuló trifón csoportosítás. Ezeket a következő részben mutatjuk be.

b) Felismerési hálózatépítés

Monofón modelleknél a felismerési hálózat építése viszonylag egyszerű, hiszen minden szó helyébe elegendő a fonetikus átíratának megfelelő monofón modellek sorozatát illeszteni. Trifón modelleknél viszont – különösen a folyamatos beszédfelismerési hálózatok esetén – azokban a csomópontokban, ahol több szót több szó követhet, nem triviális feladat a szószintről a trifón hálózati szintre történő lépés. Ekkor ugyanis attól függően kell az egyik vagy másik trifón modellt használni egy adott szó végén, hogy az őt követő szó milyen fonémával kezdődik.

A probléma szemléltetése – a példában feltételezve, hogy a beszéd szünet után indul, az első szó után viszont nem tart szünetet a beszélő („sil” jellel a szünetet jelöltük, melyet fonetikus környezet értékűnek szoktak tekinteni):

pamacs a... → sil-p+a p-a+m a-m+a m-a+cs a-cs+a cs-a+...

pamacs egy... → sil-p+a p-a+m a-m+a m-a+cs a-cs+e cs-e+gy e-gy+...

A problémának kétféle kezelése terjedt el.

1. A hálózatépítés egyszerűsítése és/vagy egyéb okokból a szószéli trifónok környezetfüggésének redukciója. Ilyenkor, mivel nem tudjuk, milyen a szókezedet esetén a bal, szóvég esetén a jobb oldali környezet, szó elején jobb oldali difón modellt, szó végén bal oldali difón modellt használunk.

A szó belseji trifón modellezés szemléltetése:

pamacs a... → p+a p-a+m a-m+a m-a+cs a-cs a...

pamacs egy... → p+a p-a+m a-m+a m-a+cs a-cs e+gy e-gy...

Ezt a megközelítést az angol nyelvű szakirodalom „word internal” trifón modellezésnek hívja, amit magyarul „szó belseji” trifón modellezésnek nevezünk. Hátránya, hogy a szóhatárokon fellépő koartikulációs jelenségeket csak impliciten, monofón szinten kezeli.

2. A korrekt, szóhatárokon is megfelelő modellek alkalmazását angolul „cross-word” trifón modellezésnek hívják, amelyet talán „szóhatárokon átívelő” trifón modellezésnek fordíthatunk. Ez a megközelítés – köszönhetően a szóhatárokon is explicit koartikulációs modellezésnek – az általános tapasztalatok szerint pontosabb beszédfelismerést tesz lehetővé, mint az előbbi módszer. A felismerési hálózatépítés részleteibe itt nem mehetünk bele, de véges automataelméleti alapokon a feladat hatékonyan megoldható.

Trifón-csoportosítási eljárások

Az „arany középut” a monofón és az általánosított trifón szemlélet között a környezetfüggő beszédhangok fonetikus környezetük alapján történő csoportosítása. Az egy csoportba eső trifónokat egyetlen (nem általánosított) trifón modellel képviseljük. Ezeket a modelleket nevezzük csoport trifón modellnek. (A továbbiakban a „csoport” jelzőt általában elhagyjuk.)

A tanító hangadatbázisok általában meghatározott körülmények között kerülnek felvételre. Ezért tipikusan nem az akusztikus (trifón, difón vagy monofón) modellekhez történik az adatbázisok igazítása, hanem fordítva. Vagyis a gyakorlatban úgy alakítjuk ki a modelleket, hogy azok egyrészt minél jobban fedjék le az adatbázis fonetikai-fonológiai gazdagságát, másrészt hogy elegendő mennyiségű mintával legyen tanítva minden egyes beszédhangmodell.

A feladat nem triviális, hiszen a két követelmény, hogy legyen minél több modellünk, ugyanakkor az egyes modelleket minél több mintával tanítsuk, ellentmond egymásnak. Az ellentét úgy oldható fel, hogy közvetlenül a modelleknek az adatbázishoz való illeszkedésének mértékét, illetve független tesztalmonzon a felismerési pontosságot mérjük és próbáljuk maximalizálni („Maximum Likelihood” és „Minimum Recognition Error” elv).

A következőkben két módszert ismertetünk (csoport) trifón modellek automatikus kialakítására. Mindkettő integrálja magában a szabályalapú és a statisztikai szemléletet.

a) Fonológiai csoportosítás

Ez a módszer onnan kapta nevét, hogy a környezetfüggő beszédhangmodellek csoportosítására nem használ fel más információt, mint az adatbázis fonológiai

statisztikáit és szabályokat. A gyakorlatban ez azt (is) jelenti, hogy a tanító beszédadatbázis két alapvető része, vagyis a hullámformák és a szöveges átiratok közül csak az utóbbi alapján történik a csoportok kialakítása.

– **Szabályrendszer:** a fonológiai szabályrendszer egyszerű szabályokból és a közöttük értelmezett hierarchiából áll. Egy szabály megadja, hogy az általa leírt bal vagy jobb oldali környezethez milyen fonémák tartoznak.

Példa szabályokra:

NAZÁLIS: *m, n...*

ALVEOLÁRIS: *d, t, n...*

VELÁRIS: *g, k...*

BILABIÁLIS: *p, b, m...*

A hierarchiát a szabályok sorrendje adja. Tehát a fenti szabályrendszert alkalmazva a *pamacs* szó első [ɔ] hangjára a „BILABIÁLIS-*a*+NAZÁLIS” trifónt kapjuk. Természetesen egy fonéma – mint környezet – több szabálnál is szerepelhet, abba a csoportba kerül, amelyik magasabb hierarchiaszinten van.

Ez a szabályalapú csoportosítási módszer a szabályrendszerbe kódolt fonológiai tudástól függően képes a trifón modellek számát hatékonyan csökkenteni azon megkülönböztető fonológiai környezetek kialakítása által, amelyeket a szabályalkotó fontosnak ítél. Azonban nem képes annak biztosítására, hogy minden így kialakított modell elegendő tanító mintához jusson.

– **Statisztika:** A tanító beszédadatbázishoz tartozó fonetikus (vagy inkább fonemikus) átiratok és a fenti szabályrendszer alapján pontos statisztika készíthető arról, hogy az egyes trifón modellekre hány beszédhangminta fog jutni. Mint említettük, ez jó, ha legalább százas nagyságrendet ér el, viszont általában semmi garancia nincs arra nézve, hogy az egyes beszédhangmodelleknél ez teljesül is.

Az újszerű javaslatunkat a következőkben foglaljuk össze. Hasonlóan a nyelvi modellezésnél az N-gramok simítására használt, eredetileg Katz által javasolt visszametszéses módszerhez (Katz 1987), azt a trifón modellt, amely nem rendelkezik elég tanító adattal, messzük vissza difónná. Továbbá ha a difón modell sem rendelkezik elég tanító adattal, akkor pedig monofónná. Hogy egy beszédhangnál mennyi tanító mintát tartunk elegendőnek, azt vegyük fel paraméternek, és úgy állítsuk be a paramétert, hogy a felismerési hibát minimalizáljuk. A megvalósított rendszerünkben a trifón és difón küszöbparaméterek külön módosíthatók, míg monofón modelleknél nem alkalmazunk küszöböt.

A fent részletezett fonológiai csoportosítás módszerével többféle adatbázissal végeztünk beszédfelismerési kísérleteket, amelyeket a következő fejezetben ismertetünk.

b) Fonetikai csoportosítás

Ez a széles körben használt megközelítés akusztikai és fonológiai információt is használ a trifón csoportok kialakításához (Young 1994). Szintén fonológiai szabályokon és statisztikán alapul, mint az előző módszer, azonban egy csoport kialakításában több (akár száz) szabály is szerepet játszhat. Az eljárást fonetikus

döntési fa alapú trifón csoportosításnak hívják, és a lényegét a következőkben foglaljuk össze.

Egy adott beszédhang trifón modelljeit (trifón csoportjait) lépésről lépésre határozzuk meg. Kezdetben az összes – adott hanghoz tartozó – trifónt egy csoportba tartozónak tekintjük, majd egy szabályt a jobb vagy bal oldali környezetre való kérdésként értelmezve, két csoportra osztjuk a trifónokat. A módszer lényege, hogy lépésenként azzal a szabállyal osztjuk ketté az arra legalkalmasabb csoportot, amelyiknek az alkalmazásával a csoport trifón modellek tanító adatbázishoz való illeszkedésének mértékét leginkább növeljük. Végeredményben így minden beszédhanghoz egy fonetikus döntési fa áll elő, amelynek levelei reprezentálják az adott beszédhanghoz tartozó trifón csoportokat. Egy (csoport) trifón modellt a csoport neve azonosít.

A szemléltetéshez tekintsük a korábbi szabályokat:

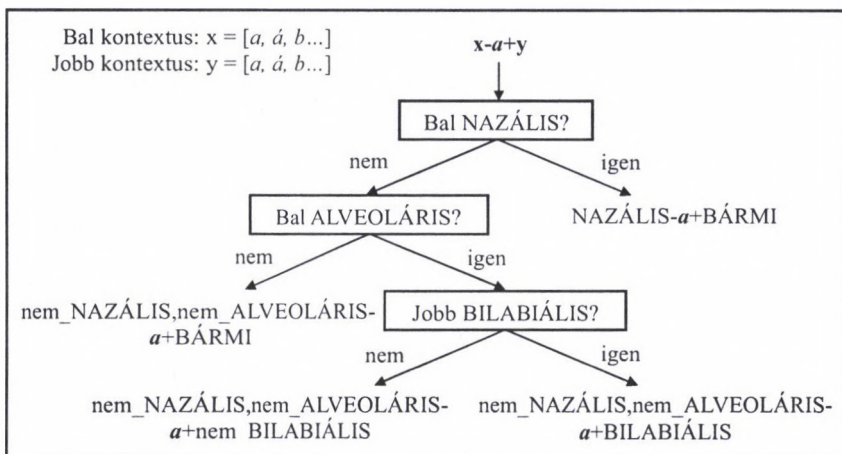
NAZÁLIS: $m, n \dots$

ALVEOLÁRIS: $d, t, n \dots$

VELÁRIS: $g, k \dots$

BILABIÁLIS: $p, b, m \dots$

Az 1. ábra egy lehetséges döntési fát mutat be az [ɔ] hangra.



1. ábra

A fonetikus csoportosítás szemléltetése

Például a *pamacs* szó első [ɔ] hangjának a (csoport) trifón modellje a szabályok és a döntési fa alapján a „nem_NAZÁLIS,nem_ALVEOLÁRIS-a+BÁRMI” lesz, míg a második [ɔ] hangja a „NAZÁLIS-a+BÁRMI” trifón csoportba kerül.

Annak biztosítására, hogy egy adott trifón csoportra elegendő tanító minta jusson, a fenti algoritmusnak csak egy minimális módosítása szükséges, mégpe-

dig annyi, hogy csak akkor engedjünk létrejönni egy új csoportot, ha egy küszöbszámot meghaladó mennyiségű tanító mintával rendelkezik.

Az algoritmus rendkívül kedvező tulajdonsága, hogy új, tetszőleges szabály hozzáadása nem ronthatja a tanító adatbázishoz való illeszkedés jóságát, valamint hogy a szabályok sorrendje tetszőleges lehet, hiszen a döntésifa-építés automatikusan, az akusztikum figyelembevételével alakítja ki a szabályhierarchiát.

Beszédfelismerési kísérletek

A statisztikai módszereken alapuló beszédfelismerési kísérleteinkben azt vizsgáltuk, hogy különféle tanítási és tesztelési konfigurációkban hogyan változik a felismerés pontossága. Minden konfigurációban három akusztikus modellezési megközelítést vizsgáltunk: a monofón modellezést, a visszametszéses csoport trifón modellezést és a döntési fa alapú csoport trifón modellezést.

Beszédatadbázisok

A tanító és tesztelő adatbázisokat a legnagyobb magyar telefonos beszédatadbázisok, az MTBA, a Besztel, a SpeechDat és a Tesztel összességéből alakítottuk ki (Vicsi 2005). Ezek az adatbázisok elsősorban olvasott beszédet, valamint kisebb arányban spontán bemondásokat is tartalmaznak. Az első három adatbázis lényegében ugyanarra a szövegkorpuszra épül, és mindegyiknek az általunk elérhető része 500 beszélőtől tartalmaz hanganyagot. A Tesztel adatbázis 100 beszélős, és jellegzetessége, hogy szándékosan nagy és természetes háttérzajban felvett bemondásokat tartalmaz. Az adatbázisokban a vonalas és mobiltelefonos felvételek összességében körülbelül ugyanolyan számban képviseltetik magukat.

Tanító halmazok

Tanítás céljára az MTBA, a Besztel és a SpeechDat adatbázis 500-400-450 beszélőjének azon felvételeit jelöltük ki, melyek nem „o”, és „z” jelzésűek, azaz nem tartalmaznak tulajdonneveket és bizonyos típusú mondatokat. A SpeechDat esetén csak egy szűkebb halmazt, a fonetikailag változatos mondatokat és szavakat használtuk.

A teljes tanító halmaz mellett annak bizonyos részhalmazait is képeztük, hogy az egyes akusztikus modellezési eljárások tanítóadatbázisméret-függését is vizsgálhassuk.

A tanító halmazok jelölése és tartalma:

M: Az MTBA fonetikailag változatos mondatai és szavai, 500 beszélő, 6 000 felvétel.

MM: Az MTBA összes tanító felvétele, 500 beszélő, 19 000 felvétel.

MM_BS: Az MTBA és a Besztel összes tanító felvétele, 900 beszélő, 39 000 felvétel.

MM_BS_SD: Az MTBA, a Besztel és a SpeechDat tanítófelvételei, 1350 beszélő, 44 000 felvétel.

Teszthalmazok

A teszthalmazokat úgy állítottuk össze, hogy ne legyen átfedés a tanító halmazban szereplő beszélőkkel. Így a tanításnál fel nem használt felvételek össze-

sen 170 beszélőtől (Besztel 100, Speechdat 50, Tesztel 20) kerültek a tesztalmazókba.

Kétféle felismerési tesztet végeztünk a tesztelő felvételek típusától függően. Az izolált szavakat, kifejezéseket tartalmazó bemondásokat (2 475 felvétel) izolált szavas felismerővel ismertettük fel, a mondatokat (2 385 felvétel) pedig folyamatos beszédfelismerővel. Számokkal, dátumokkal és egyéb, kapcsolt szavas felismeréshez illeszkedő felvételekkel nem teszteltünk.

Az **izolált szavas** halmaz azon felvételeit, amelyek lexikálisan illeszkedtek a tanító halmazhoz, azaz a tesztelő szóalakok szerepeltek a tanítás során is (pl. fonetikailag változtatott szavak), az „illeszkedő” tesztalmazba tettük. Értelemszerűen a többi felvételt (mint például a tulajdonnevek) a „nem illeszkedő” tesztalmazba tettük.

A **folyamatos** beszédfelismerési teszteknel nemcsak lexikális, hanem nyelvi illeszkedésről is beszélhetünk. Az egyik halmazba azokat a mondatokat válogattuk, amelyeket mind az akusztikus, mind a nyelvi modell tanításakor felhasználtunk, ez az „illeszkedő” halmaz (lexikális és nyelvtani szempontból is). A másik, „nem illeszkedő” tesztalmazba azok a felvételek kerültek, amelyeket sem az akusztikus, sem a nyelvi modell tanításakor nem vettünk figyelembe. Egyéb halmazt nem vizsgáltunk.

A tesztalmazók jelölése és tartalma:

I_M: Izolált szavakat, kifejezéseket tartalmazó, a tanító adatokhoz lexikálisan illeszkedő felvételek, 170 beszélő, 1 726 felvétel.

I_U: Izolált szavakat, kifejezéseket tartalmazó, a tanító adatokhoz lexikálisan nem illeszkedő felvételek, 170 beszélő, 749 felvétel.

C_M: Nyelvi és lexikális szempontból a tanításhoz illeszkedő mondatok, 170 beszélő, 1 973 felvétel.

C_U: Sem nyelvi, sem lexikális szempontból a tanításhoz nem illeszkedő mondatok, 170 beszélő, 412 felvétel.

Beszédfelismerési paraméterek, beállítások

Lényegkiemelés: A beszédfelismerési kísérleteket az alábbi paraméterek mellett végeztük. Lényegkiemelési paraméterekként a bemenő beszédjelből MFCC (Mel Frequency Cepstral Coefficients) 38 dimenziós vektorokat képeztünk csak statikus energiát használva, de egyébként dinamikus Delta és Delta-Delta értékeket is számítva. Mind tanítás, mind tesztelés során alkalmaztuk a vak csatornakiégnyelés módszerét.

Akusztikus modellek: „Balról-jobbra” struktúrájú, háromállapotú rejtett Markov-modelleket használtunk mind a monofón, mind a trifón beszédhangmodellek esetén. Állapotonként maximum 10 Gauss-függvényből álló folyamatos megfigyelési sűrűségfüggvényeket használtunk. A bal, középső és jobb trifón állapotokat a már ismertetett, kétféle – a fonológiai és fonetikai csoportosítási – eljárással csoportosítottuk. Szóhatárokon átívelő, azaz „cross-word” trifón modelleket alkalmaztunk.

Kiejtési modellek: A kiejtési modelleket, azaz a fonemikus átiratokat automatikusan állítottuk elő. Semmilyen fonetikus vagy fonológiai koartikulációt, illetve allofónikus változatokat nem jelöltünk, továbbá a hosszú és rövid mássalhangzókat sem különböztettük meg. Így – a szünetmodelleket nem számítva – összesen 39 fonológiai kategóriát használtunk. A szünetmodell háromállapotú környezetfüggetlen modell volt.

Nyelvi modellek: A folyamatos felismerésnél szó-trigram nyelvi modelleket alkalmaztunk Katz-féle visszametszéssel és Good-Turing valószínűség-újraelosztással (Church 1991). A tanító szöveg az illeszkedő tesztmondatok szövege alapján készült, minden különböző mondatot egyszer szerepeltettünk. Így az illeszkedő mondatokon PP = 40-es perplexitást, a nem illeszkedő tesztmondatokon PP = 6 230-as (nagyon magas, azaz igen kedvezőtlen) perplexitásértéket kaptunk.

Szótárak: Az izolált szavas felismeréseknél ugyanazt az 1 334 elemű szótárat használtuk az illeszkedő és nem illeszkedő felvételek esetén is. Hasonlóan, a folyamatos felismeréseknél is ugyanazt az 5 561 elemű szótárat és természetesen ugyanazt a nyelvi modellt alkalmaztuk mindkét tesztalmaz esetén. Mind a folyamatos, mind az izolált szavas tesztek esetén a teljes tesztalmazt lefedő szótárakat alkalmaztuk, így szótáron kívüli elemek kezelésére nem volt szükség.

A tanítás és tesztelés során minden konfigurációban ugyanúgy jártunk el a monofón, a visszametszéses csoport trifón és a döntési fa csoportosítású trifón modellek esetén, ideértve az akusztikus/nyelvi modell valószínűség-súlyozást és a keresési mélység beállítását is.

Tanítási eredmények

A tanítás eredményei alatt a létrejött akusztikus modelleket értjük. Míg egyállapotú modelleknél a csoport trifón modellek száma és az állapotok száma nem válik szét, a többállapotú modelleknél – mivel az állapotokat külön-külön csoportosítjuk – az állapotszám mutatja a teljes akusztikus modell részletgazdagságát (1. táblázat).

1. táblázat: A monofón, a fonológiai csoportosítású trifón és a fonetikai csoportosítású trifón akusztikai modellek állapotszáma a tanítóhalmaz függvényében

| Állapotszám | M | MM | MM_BS | MM_BS_SD |
|-------------------|------|------|-------|----------|
| Monofón | 126 | 126 | 126 | 126 |
| Fonológiai trifón | 998 | 1342 | 1706 | 1904 |
| Fonetikai trifón | 1775 | 2732 | 3952 | 4618 |

Izolált szavas beszédfelismerési eredmények

Az izolált szavas szófelismerési arányok a 2. és 3. táblázatban találhatóak. Minden esetben a fent részletezett, összesen 12 akusztikus modellt használva születtek az eredmények.

2. táblázat: Izolált szavas szófelismerési százalékarányok
illeszkedő tesztalmaz esetén

| I_M Corr. | M | MM | MM_BS | MM_BS_SD |
|-------------------|--------|--------|--------|----------|
| Monofón | 82,91% | 86,33% | 86,73% | 86,73% |
| Fonológiai trifón | 92,99% | 94,90% | 96,06% | 95,94% |
| Fonetikai trifón | 94,79% | 96,87% | 96,87% | 96,58% |

3. táblázat: Izolált szavas szófelismerési százalékarányok
nem illeszkedő tesztalmaz esetén

| I_U Corr. | M | MM | MM_BS | MM_BS_SD |
|-------------------|--------|--------|--------|----------|
| Monofón | 80,61% | 82,49% | 84,09% | 83,42% |
| Fonológiai trifón | 90,91% | 91,41% | 91,44% | 92,11% |
| Fonetikai trifón | 93,32% | 93,58% | 93,32% | 93,98% |

Látható, hogy minden oszlopban a fonetikai csoportosítás eredménye a legjobb, attól kevéssel elmarad a fonológiai csoportosítású trifón modellé és kétháromszor akkora felismerési hibával követi a monofón eredmény. Az is megfigyelhető, hogy a tanító adatbázis növekedésével általában javulnak az eredmények, de nem mindig.

A várakozásoknak megfelelően pár százalékkal elmaradnak a tanítás során nem látott hangkapcsolatokat is tartalmazó tesztalmazalon mért eredmények a másik (illeszkedő) halmaz eredményeitől. Érdekes, hogy minden konfigurációban szinte állandó 3% körüli eltérés figyelhető meg a két halmaz eredményei között.

Az egyes akusztikai modellek gyakorlati alkalmazhatósága érdekében fontos a számításigényeiket is megvizsgálni. A 4. táblázat alapján azt mondhatjuk, hogy nincs óriási különbség az egyes megközelítések között e tekintetben.

4. táblázat: Az izolált szavas gépi beszédfelismerés átlagos számításigénye
(Real-Time Factor)

| RTF | M | MM | MM_BS | MM_BS_SD |
|-------------------|-------|-------|-------|----------|
| Monofón | 0,066 | 0,067 | 0,067 | 0,067 |
| Fonológiai trifón | 0,089 | 0,105 | 0,110 | 0,120 |
| Fonetikai trifón | 0,077 | 0,095 | 0,127 | 0,130 |

Folyamatos beszédfelismerési eredmények

A folyamatos beszédfelismerésnél lényeges újdonság, hogy nyelvi modell alkalmazására is szükség van, melynek minősége nagymértékben befolyásolja a felismerési eredményt. A nyelvi modell használatának célja, hogy csökkentse az akusztikus modellre háruló döntés nehézségét azért, hogy valószínűségi becs-

lést ad a felismerési szósorozatra – pusztán szövegstatistikai alapon. A nyelvi modell PP értéke szemléletesen azt mutatja meg, hogy egy szó után átlagosan hány egyformán legvalószínűbb szó következhet.

Az 5. és 6. táblázatban láthatók a folyamatos felismerési eredmények. A jóval összetettebb feladat ellenére az illeszkedő tesztfelvételek esetén hasonló felismerési arányokat kaptunk, mint az izolált szavas teszteknel. Észrevehető viszont, hogy a tanító adatbázis méretének növelésével a monofón felismerési eredmény romlott, és a trifón eredmények is alig javultak.

5. táblázat: A folyamatos beszédfelismerési tesztek szófelismerési százalékarányai **illeszkedő** tesztalmaz esetén

| C_M Corr. | M | MM | MM_BS | MM_BS_SD |
|-------------------|--------|--------|--------|----------|
| Monofón | 85,34% | 80,19% | 79,87% | 80,74% |
| Fonológiai trifón | 92,72% | 91,43% | 92,15% | 92,84% |
| Fonetikai trifón | 94,13% | 93,82% | 94,22% | 94,47% |

6. táblázat: A folyamatos beszédfelismerési tesztek szófelismerési százalékarányai **nem illeszkedő** tesztalmaz esetén

| C_U Corr. | M | MM | MM_BS | MM_BS_SD |
|-------------------|--------|--------|--------|----------|
| Monofón | 29,22% | 25,40% | 24,72% | 25,58% |
| Fonológiai trifón | 52,16% | 50,30% | 53,63% | 54,88% |
| Fonetikai trifón | 61,95% | 61,27% | 64,24% | 64,42% |

Drasztikus teljesítményromlást láthatunk viszont a nem illeszkedő felvételek esetén. Itt mutatkozik meg a nyelvi modellhez való illeszkedés jelentősége. A nagy különbség mindazonáltal jól indokolható a két tesztalmazon mért gyökeresen eltérő perplexitásértékekkel (PP).

Tanulmányunknak talán legfontosabb megfigyelése a 6. táblázatból olvasható ki. Jól látható, hogy a nem illeszkedő tesztalmazon a monofón modellek folyamatos felismerési eredménye, azaz a felismerési általánosító képessége összehasonlíthatatlanul gyengébb a trifón modellekénél. Míg az egyéb teszteken a monofón felismerési hiba volt fele-harmada a trifón hibának, addig itt a felismerési **arány** feleződik-harmadolódik meg a trifón esethez képest.

Vagyis a monofón modellek felismerési aránya a korábbi hibaarányainak szintjére süllyedt (kb. minden 4. szó helyes csak!), míg a jobban teljesítő fonetikai csoportosítású trifón modellnél a szavak csaknem kétharmadát továbbra is helyesen ismertük fel.

Ahogy a 7. táblázat mutatja, hogy a trifón modellek sokkal jobb folyamatos felismerési pontossága nem jár együtt hasonló mértékű számítási igény-növekedéssel. Sőt még kisebbek a különbségek, mint az izolált szavas felismerésnél.

7. táblázat: Az izolált szavas gépi beszédfelismerés átlagos számításigénye (Real-Time Factor)

| RTF | M | MM | MM_BS | MM_BS_SD |
|-------------------|-------|-------|-------|----------|
| Monofón | 0,540 | 0,607 | 0,607 | 0,607 |
| Fonológiai trifón | 0,770 | 0,840 | 0,890 | 0,900 |
| Fonetikai trifón | 0,570 | 0,690 | 0,780 | 0,850 |

Összefoglalás

Bemutattunk egy új módszert környezetfüggő beszédhangmodellek kialakítására. Az eljárást fonológiai trifón csoportosításnak neveztük, mert pusztán fonológiai szintű szabályrendszer és statisztikák alapján történik a környezetfüggő beszédhangmodellek kialakítása. A javasolt megoldás ötvözi a szabályalapú és a statisztikai szemléletet.

Módszerünket összevetettük a környezetfüggetlen és a széles körben használt fonetikai döntési fa alapú trifón csoportosítási eljárással. Számos beszédfelismerési kísérletet végeztünk nagy mennyiségű és több mint ezer beszélőtől származó tanító és tesztfelvétellel. Izolált szavas és folyamatos beszédfelismerési tesztet egyaránt végeztünk különféle módon illeszkedő tesztalmazokkal.

Tapasztalataink szerint az általunk javasolt megoldással ugyan nem sikerült elérni a fonetikai csoportosítással kapott modellek felismerési pontosságát, azonban megközelítettük azokét annak ellenére, hogy körülbelül feleannyi állapotot használtunk. Ugyanakkor a monofón modellezéshez képest igen kedvezően teljesített a javasolt módszer.

Végkövetkeztetésünk, hogy a környezetfüggő beszédhangmodellek minden körülmények között lényegesen jobb felismerési hatásfokot látszanak biztosítani, mint a környezetfüggetlenek, folyamatos beszédfelismerésnél használatuk – általánosító képességük miatt – szinte „kötelező”.

Irodalom

- Church, Kenneth W. – Gale, William A. 1991. A comparison of the enhanced good-turing and deleted estimation methods for estimating probabilities of English bigrams. *Computer Speech and Language* 5. 19–54.
- Jurafsky, Dan – Ward, Wayne – Jianping, Zhang – Herold, Keith – Xiuyang, Yu – Sen, Zhang 2001. What kind of pronunciation variation is hard for triphones to model? In *Proceedings of ICASSP-01 Vol I*. Salt Lake City, Utah, 577–580.
- Katz, Slava M. 1987. Estimation of probabilities from sparse data for the language model component of a speech recognizer. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 35/3. 400–401.
- Olasz Gábor 2003. Az artikuláció akusztikai vetülete – a hangsebészet elmélete és gyakorlata. In Hunyadi László (szerk.): *Kísérleti fonetika, laboratóriumi fonológia a gyakorlatban*. Debreceni Egyetem Kossuth Egyetemi Kiadója, Debrecen, 241–254.

- Siptár Péter 1995. *Fonológia*. Egyetemi jegyzet. MTA Nyelvtudományi Intézet, Budapest.
- Vicsi Klára et al. 2005. Hungarian speech databases. <http://alpha.ttt.bme.hu/speech/databases.php>.
- Young, Steve – Kershaw, Dan – Odell, Julian – Ollason, Dave – Valtchev, Valtcho – Woodland, Phil 2000. The HTK Book (Version 3.0). <http://htk.eng.cam.ac.uk/>.
- Young, Steve – Odell, Julian – Woodland, Phil 1994. Tree-based state tying for high accuracy acoustic modelling. In *Proceedings of ARPA Workshop of Human Language Technology*. 307–312.

„NYELVBOTLÁS”-KORPUSZ

3. rész

Az első magyar valós idejű, jegyzeteléses technikával gyűjtött megakadásjelenség-korpusz első részét 5139 tétellel 2004-ben publikáltuk (*Beszéd kutatás 2004.* 19–186); második részét 761 tétellel pedig 2005-ben (*Beszéd kutatás 2005.* 145–173). A gyűjtés és az osztályozás azonban ezzel nem maradt abba, és a korpusz további részeit folyamatosan közzé kívánjuk tenni – e helyt újabb 388 megakadás olvasható. A kategorizálás elméleti háttérét részletesen bemutatja Gósy Mária bevezető tanulmánya: A spontán magyar beszéd megakadásainak hallás alapú gyűjteménye (*Beszéd kutatás 2004.* 6–18).

Az adatokat kategóriánként közöljük, azon belül pedig ábécésorrendben. A kategóriák eredeti sorszámozását (vö. *Beszéd kutatás 2004.*) megtartottuk, akkor is, ha az adott kategóriába ezúttal nem került adat, és így kimaradtak sorszámok. A „Megakadás” oszlopban található a megakadást tartalmazó nyelvi adatok, a „Szándékolt közlés” oszlopban pedig a beszélő eredeti tervének megfelelő nyelvi formák – amennyiben a beszélő a téves megnyilatkozásban nem javított, és így abban az eredeti szándék nem jelenik meg. A □ jellel a szünetet jelöltük. A három pont (...) arra utal, hogy az adott közlés további részletei nem ismertek. A × jelet a kontaminációknál alkalmaztuk, a vegyülésben szerepet játszott, eredeti formák szemléltetésére.

2. Morfológiai, szintaktikai hiba**Megakadás**

A tanárok bementek a teremre.
 Egy súlyos nyomokat akart eltüntetni az illető.
 Együtt jár az alkoholnak vagy a drogfogyasztással.
 Először egy egért ölt meg.
 Hagytad este a családnak a tortából?
 Hálósákokat legfőljebb sátorban vagy a szabad égen...
 Nem kell kényszerítenie magukat...

Szándékolt közlés

terembe
 egy súlyos nyomot
 alkohol- vagy drogfogyasztással
 egeret
 hagytaál
 ég alatt
 kényszeríteniük

2. a) Újraindítás nélküli morfológiai változtatás**Megakadás**

Aki a technikus ☐ si posztra vágyik, az egy teljesen más történet.
 Akkor is, ha négy perccel beszélsz egy órán ☐ ban vagy negyvenet.
 Cél az, hogy a fiatalok ☐ at tájékoztassuk.
 Egy lány felsőfokú német nyelvvizsgából ☐ ő val oktatást vállal.
 Érettségi tétel ☐ ő ☐ ek...
 Igazából csak le kellene mennem hozzá ide az első emeleten ☐ re, és beszélnem vele.
 Jártam beszédórara, de az már itt volt a Danubiuson ☐ ban.
 Kiterjed más városokra és kerületekben ő re is.
 X a minimum- ☐ -málbért kapja.

Szándékolt közlés

tételek

minimálbért

2. b) Ismétlés toldalékjavítással**Megakadás**

A 80-as évek vége, a 80-as évek végén...
 A diabétesznek többféle okai is lehetnek, oka is lehet.
 A számonkérésben a számonkérés nagyon érdekes.
 Az egyesület a családokat karolták föl ő karolta föl.
 Az is elmondható, hogy itt már tulajdonképpen jól működő tanárokról vannak ☐ van szó.
 Azt mondta, hogy ő ez, hogy őt ez érdekli.
 Én meg beleegyeztem abban ☐ abba...
 Ez a bizalom ☐ bizalmi viszony sérült.
 Itt számos dolgot ☐ őö dolgot megtanulnak a gyerekek.

Megakadás

Hat hetet, hat hét lesz az első negyedév.

Kaptatok olyan üzenetet már, ami nem a témának □ témáról szolt, hanem nektek személyesen?

Keresi a szavat □ szót □ szavakat.

Már szinte le vagytok késve arra □ arról, hogy ilyet csináljatok.

Még egy dolgot szokott □ szoktak emlegetni a tizeddel kapcsolatban.

Mindjárt megtudjuk, mik azok az öt számok □ mi az az öt szám.

Nagyon sok családdal ő családnál teszek látogatást.

Nagyon szeretem □ szerettem volna ezt a számítógépet.

Ne hamarkodd el ez □ ezt!

Nem kapják meg a gyógyszere gyógyszert.

Nem kifejezetten irodalom □ irodalmi érdeklődésű.

Nem tudom, mennyivel ő mennyire fogyasztok egészséges ételeket.

Nem vagyok egy nagy türelmes □ egy nagyon türelmes ember.

Soha életemben, ez a □ ezt a Petra nevében is mondhatom, sose gondolkoztunk erről.

Sokféle gondolat cserélt □ sokféle gondolatot cseréltek ki egymás közt.

Tehát tudni □ tudjuk ilyenkor, hogy mi következik.

Tehát valószínűleg ezt □ ez lenne a kenyérkereső dolog most nekem.

3. Kontamináció

Megakadás

A firte a fiatal Szirtes Ádám.

– Az alkoholtól fölébred vagy elalszik az ember?

– Hát, inkább elálmosodik tőle. – És ha már eleve fáradt, akkor fölalszik? Öő elalszik?

Az ember egya megáll □ na agya □ vagy esze.

Az ő nérvetei □ nézetei...

Deák és Kossuth, akik szögesen nem értettek egyet egymással.

Déry elkezdí átértél átértelmezni és átértékelni...

Ennek a gyógyszernek a hatását csökkenti a kea □ na, a tea, kávé, tej.

Folyadékállóval kell feltölteni.

Ha természetesen módon természetes módon viselkednél...

Hívnak létre...

Hogyan szoktunk gondolkozni a keleti országokra és az onnan jövő munkásokra?

Jó volna vehetni jegyet.

Szándékolt közlés

fölebred × elalszik

Az ember esze megáll.

nézetei × érvei

akik szöges ellentétben álltak × nem értettek egyet

fagyálló folyadékkal

természetesen × természetes módon

hoznak létre × hívnak életre gondolkozni a keleti országokról × gondolni a keleti országokra

jó volna venni × jó volna, ha lehetne venni

Megakadás

Mondtam neki, hogy ez nem erről működik.

Nagyon kártó lehet.

Nagyon röviden próbálok lenni.

Nem csukadt be rendetted ☐ melletted rendesen az ajtót.

Sokkal tovább ☐ tovább tart.

Tangolni jobb.

Te ne játsszad itt az aszad ☐ eszed.

Szándékolt közlés

erről szól × így működik

káros × ártó

röviden fogok beszélni ×

rövid próbálok lenni

4. Téves szótalálás**Megakadás**

A Csonginak legót vettünk karácsonyra, izé húsvétra.

A fején ☐ vagyis a kezén kell hordani a sisakot?

A hajam szélnek áll.

A minap volt ☐ lesz egy ilyen valóságshow, ami pont ezzel fog foglalkozni.

A zarándokok háromlatai lelassultak.

Addig ütik a vasalót, míg meleg.

Adok valamilyen vonal ☐ sorvezetőt.

Az alábbiakban továbbiakban fogok beszélni...

Az édesapa megtanul anyává, apává válni...

Bármennyire is betanítjuk, betartjuk a normál eloszlás...

Bekönyörgök beköszönök Leventéhez.

Boldog karácsonyt, dehogy, húsvétot!

Elfogadták elfoglalták Pannóniát.

Elképesztő ☐ ööö elképzelhető ez is.

Emlékszel, amit feltett a kisugárázóra ☐ kivetítőre.

És nem csurog a nyelved ☐ nyálad?

És ott vannak azok a pártok, amelyek jól sikerül ☐ jól szerepeltek.

És utána vagy mondja a korrekt ☐ korrigálást vagy nem.

Ezekből a labdaszerződésekből meg tudnak indulni.

Ezt a hagymát meg kéne pirítani.

Gyerekeknek tett ígéretet nem lehet megszakni.

Ha a C-ből sikerül az egyik, akkor úgyis megkapom a ☐ diplomát.

Szándékolt közlés

égnek

áramlatai

vasat

labdaszerzésekből

szallonnát

megszegni

nyelvvizsgabizonyítványt

| Megakadás | Szándékolt közlés |
|---|----------------------------|
| Hogy például a hivatali zsarnok őö hivatali zsar- gon annyira bonyolult. | |
| Hozd el a számítógépet ööö számológépet is! | |
| Hú, így meg lehet fogni a lábánál keresztül. | a lábánál fogva |
| Húsvétkor a bekötött utakon is közlekedési ellen- őrzés lesz. | bekötőutakon |
| Így tanultam meg a nyelveket, neveket is. | |
| Innen terül terítik szét... | |
| Itt vagyok tőled fél méter ööö órára! | |
| Javaslatokkal állnak szembe □ elő. | |
| Jó a hajad □ fülbevalód! | |
| Kistelepülések postáinak megszűntetéséről volt szó. | megszüntetéséről |
| ...koncentrátumos körök... | koncentrikus |
| Köszönöm a találgatást □ vagyis a tájékoztatást! | |
| Lángol □ ég a cigim. | |
| Lehet, hogy a többit is rá fogom ragasztani, aka- rom mondani, rakni. | |
| Már a még az adatfelvétel előtt... | |
| Megjelöljük a harmadikat, akarom mondani meg- nézzük a harmadik kérdést. | |
| ...mert berúglak... | megrúglak |
| Mióta átrakták az órákat, mindig álmos vagyok. | átállították |
| Most ifjabb Plinius: A Venus kitörését vesszük. | Vezűv |
| Nagy a korlát □ öö bocsánat nagy a szakadék a két képzési forma között. | |
| Nagy ellenállást fog ellenem tenni. | ellenállást fog tanúsítani |
| Nagy érdeklődést mutat a kortárs novella felé ö íránt. | |
| Nem fontos itt időt vesztegetni. | vesztegetni |
| Nevezd már be, hogy ki volt az! | nevezd már meg |
| Olyan a beszéd típusa □ vagy beszéd stílusa. | |
| Orvosokkal is kidolgoztuk kitöltöttük a kérdőívet. | |
| Sok benne a vitamin, mert van benne citromsava- nyú sav. | citromsav |
| Szerencse, hogy vastag a szőr a bundáján □ a nyakán. | |
| Takard meg ezt a kislányt! | takard be |
| Túlságosan nagy a léc. | magas |
| Úgy is meg kell □ vagy mi az, ki kell mosni a függönyt. | |

| Megakadás | Szándékolt közlés |
|---|-------------------|
| Valaki kibontotta tegnap a biztonsági betéteket <input type="checkbox"/> vagy mik azok, tisztasági betéteket. | |
| Választhatja-e magának a pedagógus az osztály-főnököt? | diák |
| Valószínűleg vele hasonlóan tehetséges emberek voltak. | hozzá hasonlóan |
| Van benne elszánás. | elszántság |
| Várni kell a tüdőszűréssel, mert elromlott a rög-töngép. | röntgengép |

4. a) Téves szókezdés

Megakadás

A cseheknek van egy emigráns kormánya Mosz ☐ Londonban.
 A dolog másik ré ☐ másik oldala nehezebben megy.
 A gyermek ☐ a fiú úgy búcsúztatja...
 A joghurtot a Sm ☐ a Tescóban vettem.
 A rá ☐ a tévzés pedig egy istenadta plusz nekem.
 A szülők választanak egy nekik megte ☐ tetsző nevet.
 ...az európai un út felé tovább menni.
 Azt hiszem, a harma ő negyediken van.
 És vágtunk vörös ööő lilahagymát is.
 Ezért alakul át arr ☐ abba az irányba.
 Ezt meg tedd bele egy lá ☐ dobozba.
 Hát erről így köny ☐ külön nem olvastam.
 Itt egy ilyen megjegyzés len ☐ volna helyénvaló.
 Kapcsold már ki a mi mosógépet!
 Köszönöm, most nem kérek. Az előbb ebé ☐ vacsoráztam.
 Majd gon ☐ megpróbálnak segíteni.
 Nem per ☐ nem feltétlenül.
 Nyilvánvaló előfelté ☐ előfeltevései vannak.
 Prolaktin, ennek a szintje folyamatosan term emelkedik.
 Történelmi drámáihoz nagyon sokat tanulmányozta a történelmi műveket, például a sze ☐ a Galileihez a Galilei életét.

4. b) Malapropizmus

| Megakadás | Szándékolt közlés |
|---|-------------------|
| Hihetetlen, hogy milyen a faunája annak az or-szágnak, gyönyörű virágok vannak ott. | flórája |

5. „Nyelvem hegyén van”

Megakadás

Szándékolt közlés

A magyarítás pedig az a $\square \ddot{o} \square$ nem jut eszembe a szó \square nem folyamat, hanem \square mondjuk törekvés...

És a plázában a kínai meg görög étterem mellett lenne romániai □ román □ na, hogy kell mondani, roma étterem?

És itt akkor aaz éle □ öö vesszőfoga...

méregfoga

És mondta, hogy egyik □ egyik □ egyszer, amikor kint járt...

Ez lényegében ugyanaz, mint amit már elmondott
aa □ na, segíts már, Vera.

Frizuá □ frizuá □ haj...

frizura

Lezuhant egy magyar heggy ☐ hegye ☐ sziklamá-
szó.

Magasabb s ☐ ss ☐ sa ☐ szárú csizmát akarok.
Mindig én felejte ☐ felem ☐ feleltem kiskorom-
ban történelemből.

...modi modu módosít...

modulál

Nem értem, hogy mit jelent, hogy szó szerint be-
viharzott az érte ☐ étterembe ☐ na, a terembe.

Ott van a kenyérszete ő kenyérszelető, na, kenyérszeletelő.

Rég dolgozunk együtt, tudunk ☐ tudjuk ☐ ismerjük egymás gyengéit.

Zöldséges rizzsel ☐ vagyis sajttal ☐ ízé.

tojásos rizzsel

6. Változtatás

Megakadás

Azok pedig olyan dolgok, amik evidens \square evidenciaszámba mennek.

Ez a része az egésznek nincs \square nem áll hozzám közel.

7. Újraindítás

Megakadás

A könyv négy nagy ré részre osztja ezt a jelenséget.

Az a baj, hogy e \square ezek is...

Az első monda \square mondata bevezető...

Ettől lesz olyan a műsor, amilyen, de a me \square a menete az gyakorlatilag ennyi.

Ez az alak te teljesen összevissza beszél.

Ezek az agyhu agyhullámok...

Megakadás

Különböző információk érik sz szünetben az embert.

Mikor írnak be jobban az emberek? Amikor mérgesek, vagy amikor helyesl helyeslik azt, amit mondotok?

Ő a mai kor Lata ☐ Latabárja.

Ti pi ☐ picit bepánikoltatok, nem?

Tudnak válogatni a le a le a lehetőségek közül.

Tudnak válogatni az életpálya köze közepén.

7. a) Újraindítás változtatással**Megakadás**

Amúgy ő normá ☐ viszonylag normális körülmények között élt Szárszón. Bűszke a tudá a táncstudására.

De azért ☐ de a végén egész jól összejött.

Ez tulajdonképpen erre az időre ☐ bizonyos esetben erre az időre alakult ki.

Hát nagyon jó ☐ nagyon gyönyörű volt minden.

Lehetett érezni egyfajta ☐ ezeknek a tanároknak az esetében egyfajta megszo kást.

...másokra milyen benyomást mások mit gondolnak rólam.

Mennyire sűrűn mennyire gyakran...

Mi pont azt próbáljuk meg ☐ pont azt a rést próbáljuk megtalálni, ami érdekl az embereket.

Mondom, nem tudom pon ☐ nem értem, miről szólt pontosan.

Nem hisze ☐ hiszitek el.

Több ezer forintos havi ré havi törlesztőrészlet...

8. Perszeveráció**Megakadás****Szándékolt közlés**

A lelki segítségen túl nagy segítség lenne ő nagy szükség lenne anyagiakra is.

A magyart tanítom csak, ez egy magy ő magas óraszámú tárgy.

A nagybátyám cukorbeteg, azért nem ehet kenye ret, de cukrot igen ☐ őő zsemlét igen.

A pine-dining [pajn dajniņ] kárátén belül van erre lehetőség.

...alaposan megrágva, nyálan ő nyállal elkever ve...

Az ágyunk úgy van, hogy a fejrész az ágynál ☐ őő a falnál.

Egy kötegbe röktam ☐ raktam az iratokat.

a pine-dining keretén belül

Megakadás

Én már sírással küsködtem.
 Én személyesen kérek elnézést.
 Érkezik hozzánk valanki.
 És akkor jött a macska, és a macskával □ őő a
 mancsával belekapaszkodott a szűnyoghálóba.
 ...fluidok és szerves savok ő savak...
 ...hanem vannak kérdések, és az erre adott kér-
 dések meg az arra adott kérdések korrelálnak.
 Hatvanéhes volt a mesterszakács.
 Ismeretlen eledetű testi tünetek...
 Jó, de csak egy pellanatra.
 Lehet, hogy elmegy arra, és ott karing.
 Mi az, amiben fejlesztésre so szorol?
 Mindenki a saját magához legközelebb álló témá-
 hoz beszél.
 Ne köszöröljük köszörüljük a nyelvünket!
 ...sörösdobos...
 ...száz karton szigaretta...
 Tudnak ezek jók is lenne □ lenni.
 Utalványokat vagy pel pedig ruhaneműt...

Szándékolt közlés

küszködtem
 személyesen
 valaki

 az erre adott válaszok meg
 az arra adott válaszok
 hatvanéves
 eredetű
 pillanatra
 kering

 témáról

 sörösdoboz
 cigaretta

9. Anticipáció**Megakadás**

A gyerekek kiszipolyá □ kiszipolyozása...
 A kilencéves évek óta vagy kilencvenes.
 A lehajtott fejet a sab □ szablya sem vágja le...
 A perenyi parenycát kóstold meg!
 A savazatok között kisorsoljuk...
 A tanulmányt inkább szaját szavaimmal idézem.
 A Törökvész úton végén...
 A Tuszán □ Tiszántúlon élőkre is vonatkozik.
 ...a vi vágott virág...
 Abszolút lelkesen üdvözölte az őszirózsás forrad-
 dalmat.
 Ajmát majsolt.
 Almás és sárgabarackos karcúsúzó turmi[ks].
 Amikor olvasunk egy szöveget □ olvasunk...
 Amit nyi nyi ritkán szoktunk nyíltan kérdezni...
 Ast is lefényképezem!
 Az oktatás második éves szeregszemle □ sereg-
 szemléjét...

Szándékolt közlés

szavazatok
 saját
 út

 őszirózsás

 almás
 karcúsító

 azt

| Megakadás | Szándékolt közlés |
|---|-------------------|
| Az unokatesóm szerzett egy porogramot... | programot |
| Azért, mert snem □ nem volt benne se x, se y. | |
| Aztán neked kellett visszaállnod a régi kevéskvágásba kerékvágásba. | |
| ...bevételizni a gyógyszert... | bevételezni |
| De mi van azokkal, akik mondjuk uto □ gimnázium utolsó évében vannak? | |
| Én nagyon sokat köszönhetek tőlük, nagyon sokat tanultam tőlük. | nekik |
| Erik szülőfű □ szőlőfürtöt hozott. | |
| Ez ellen már nem lehet szélhalo □ szélmalomharcot vívni. | |
| Főleg úgy, hogy a sós □ a hús egy kicsit sós volt. | |
| Ha i á elhagyjuk az ideát... | |
| Hagyd máj a tokaji borozót! | már |
| Három tallár is volt a kiránduláson. | tanár |
| ...hirdetett meg a gimnázium, és ezzel rögtön élt a ő szóval hirdetett meg a minisztérium, és ezzel rögtön élt a gimnázium. | |
| ...hogy kicsi □ a korábbi adókat kicsikarja. | |
| Jó alaposan mogol fogat. | mosol |
| Két nyelve □ éve nyelvvizsgázott. | |
| Legyen aratereje! | akaratereje |
| Lelöljük meg a D választ! | jelöljük |
| Majd ha megolvastam az e-mailt, elolvasom azt is ő megírtam. | |
| Meghatározható a ra haszonracionalizáció. | |
| Mielőtt ölre voltak őö mielőtt ölre mentek volna. | |
| Mögöttek is jöttek emberek. | mögöttem |
| Náluk van egy konyha a konyhában □ őö egy hűtő a konyhában. | |
| Ne csoroljuk a családfát! | soroljuk |
| Nem a szab □ kábítószert használata... | |
| Nem kell becsukni az ablakot, csak a függöny □ függöny mögé tenni. | |
| Pártai Lucia szerint ragyogós, napos □ na, ragyogó, napos időnk lesz ma. | |
| ...poloszkafészkes paraván... | poloszkafészkes |
| Szeretném csinálni látni, hogy rendesen csinálod. | |
| Sztili sztili stilisztikusok között... | |
| Szülőknék, pedagógusoknak és izn □ intézményvezetőknék... | |

Megakadás

Tudod, mit ettem Lipcsében? Megyve ☐ medvehagymát.

Utálom az olyan embert, aki hátulról mosolyog, és leszúr.

Szándékolt közlés

mosolyog, és hátulról leszúr

10. Metatézis

Megakadás

A Millenárison nagyon sok porgram van.
Amit János megtanult, Jancsi se felejtí el.

Azonk ☐ azoknak a tulajdona...

...beszédprodóziái prozódiai jelenségek...

...cirkecsomb...

Egy is van otthon ☐ egy otthon is van.

Elmondta nekem elmondtam neked...

És akkor egyszer csak kihívták az iskolát a rendőrs őő a rendőrséget az iskolába.

Fel lehet tenni több képes it. ☐ Ááá ☐ mit mondtam? ☐ Több képet is.

Így kívülről így be ☐ így belülről kinézel...

Kengyel gyalogkakukk futó ☐ naa, kengyelfutó gyalogkakukk.

Láttam Sábri Fanyit.

Még le tudja se lapít ☐ még le se tudja lapítani.

Megvan a szára széledve.

Nem adta oda a gyereket a könyvnek.

Összefut a szál a nyámban.

Öt kaja szatyorral...

Sorba és börbe úszott a konyha.

...tavaszi nékposzás...

Szándékolt közlés

program

Amit Jancsi megtanult, János se felejtí el.

csirkecomb

Fábri Sanyit

a széle száradva

a gyerekeknek a könyvet

a nyál a számban

öt szatyor kajával

sörbe(n) és borba(n)

népszokás

10. a) Az egyik felcserélt elem nem jelenik meg (zéró elem, vagy a beszélő a kimondása előtt megáll)

Megakadás

Beraktam a videót.

Kaptam e-mailt Oxanából Csehor ☐ Oxanától Csehországból.

Most már tudom, mi az, és hogy nem tudok tel-lene enni ☐ őő ellene tenni semmit.

Szándékolt közlés

a kazettát a videóba

11. Egyszerű nyelvbotlás

11. a) Betoldás

Megakadás

...ananaszslé...

Krösz!

Legalább nem lesz belőle portfógó ☐ öö porfogó.

...trapplista sajt...

Szándékolt közlés

ananaszslé

Kösz!

trappista

11. b) Csere

Megakadás

12 ezer embertől földöntúli ☐ földöntúli erőt kap-
tok?

A Csabu jobban rád van akaszkodva.

A héten biztosan fogok rottózni.

A jelek se ☐ szerint nem.

A kosiban van.

Árokba hajtott egy kamilon.

Egy cséfe kávé.

Ez a legeltermedtebb...

Fel lehet fejezni ☐ fel lehet fedezni...

Feled más volt, tényleg!

Gazdagok és szemények öö szegények közötti...

Nem szeretsz komprám ☐ kompromisszumokat
kötni.

...serdülő fiú rajta...

Tipikus tánciskolás gatyá.

Szándékolt közlés

Csaba

lottózni

kocsiban

kami[ʃ]on

csésze

legelterjedtebb

veled

rajza

tánciskolás

11. c) Kiesés

Megakadás

A fény hol előrefele, hol hátfele mozgott ☐ hátra-
fele.Amennyire ez kikövetkeztető ☐ kikövetkeztethe-
tő Illyés feljegyzéseiből...

Annak a képviselői szétspic szétspriccelnek.

Aztán a sorsdöntő találkóra sietett ☐ találkozóra.

...és otorozta...

Fertőzésekkel szemben kiszolgáltatókká válunk.

Határan öö határozottan jó volt az előadás.

Ilyen katoszerűe ☐ katonaszerűségek...

Szándékolt közlés

ostorozta

kiszolgáltatottakká

Megakadás**Szándékolt közlés**

Láttad a Megatár ☐ Megasztárban azt a 19 éves
főállású anyát?

Nincs más választom választásom.

...szexuális molesztásával ☐ molesztálásával vá-
dolt...

12. Többféle osztályozás**A magánhangzó-harmónia megsértése vagy túlszabályozása; egyszerű
nyelvbtlás****Megakadás****Szándékolt közlés**

A tenispályánél találkozunk.

Két lépést kell tenni, és kifülled.

tenispályánál

kifulladás

Anticipáció; egyszerű nyelvbtelés**Megakadás****Szándékolt közlés**

El van neveztgetve.

Ez egy anci ő anticipáció volt.

Ez nagy ostobá ostobaság.

Ezzel mágyarárazható...

Fed ☐ feledhetetlen este a mai!

Kábítószer ájtása mi^[j]jatt indítottak vizsgálatot.

Mik voltak az ér ellenérvek?

Tizenegy órába ☐ irányába elindulok.

nevezve

magyarázható

árusítása

óra irányába

Anticipáció; kontamináció**Megakadás****Szándékolt közlés**

És aztán, amikor csinget a mikró.

csilingel × csenget

Anticipáció; metatézis**Megakadás****Szándékolt közlés**

Alonso rejartolja...

Erről a könyvről már négy ☐ erről a versről már
négy könyvet írtak.

Mert ha beng ☐ belegondol az ember...

Nagyon erősen demozari ☐ demoralizálja.

Teljes felbárodozást okozott ez a szörnyű eset.

lerajtolja

felháborodást

Anticipáció; metatézis; perszeveráció**Megakadás**

Bartók forogja a □ sírna a □ forogna a sírjában.
Lekankadt a lóka ő mi? Lekókadt a kankalin.

Anticipáció; morfológiai, szintaktikai hiba; perszeveráció**Megakadás**

Jönnek megnézni a teleket.

Szándékolt közlés

telket

Anticipáció; perszeveráció**Megakadás**

A spontán mutáció, tehát degeneráció létén, vagy-
is lévén jött létre ez a kutyafaj.

A többiek arcára kiült a reges rémü □ jeges rémü-
let.

Az elkészítése sze sem egyszerű.

Azt mondták, ez a félórás bet beszélgetés sokat
számított.

Ebben a kontek öö kontaktusban létezhetem.

Egyetlen ember akart engem megákár □ akadá-
lyozni...

Ellentétben azokkal a diplomázott emberek □
emberekkel, akik gyengén szerepeltek...

...elsősorban diszkódrag drogoknak...

Erre egyébként nem kértem ki □ nem tértem ki.

Ez is a Kerespesi temetőben történt.

Ez nem lehetséges a szerelem beteljesel □ betel-
jesülése nélkül.

Ezt a versenyt elvesztette ugyan, de remélhetőleg
ez a helyzet □ helyzet nem fog állandósulni.

Gyúrjunk a gyomjfekek[j]re!

...infrastruktrulá na infrastrukturalis...

Saját gyerekkorájól szóló fejezetek...

Szóbeli, illetve írásbeli figyelmeztetésben is ré-
szesítette...

Szóval van Észak-Olaszország és Dél-Osza □
Olaszország.

Valahogy elkanyalodtunk...

Szándékolt közlés

Kerespesi

gyomorfekeklyre

gyerekkoráról

figyelmeztetésben

elkanyarodtunk

Anticipáció; perszeveráció; téves szótalálás

Megakadás

Amíg ki nem hal hül az étel.

Anticipáció; téves szótalálás

Megakadás

Szándékolt közlés

Akkora kincsnek vagyunk a tuda □ a tulajdono-
sai...

Civil szervezet □ szervezet tagja vagyok.

Elő ő eleinte is ők jöttek.

Ezek függnék a reflektáló művész műveltségétől. nyelvész

Hú, ezek kifejezett öö nagyon jó kifejezések!

Nagy tulajdonságot ő jelentőséget tulajdonítottam
neki.

Rákutcában járunk.

zsákutcában

Egyszerű nyelvbottlás; kontamináció

Megakadás

Na, védlőbeszéd, hú, szóval vádbeszéd és védőbeszéd?

Egyszerű nyelvbottlás; metatézis

Megakadás

Szándékolt közlés

Kés, villa, kalány.

kanál

Egyszerű nyelvbottlás; metatézis; malapropizmus

Megakadás

Szándékolt közlés

Az apukulturás gyógyítás nem csodamódszer.

akupunktúrás

Egyszerű nyelvbottlás; morfológiai, szintaktikai hiba

Megakadás

Szándékolt közlés

A tartalom az egy sokkal összetett dolog.

összetettebb

Amikor már érezik érezzük...

Ez egy olyan idegpálya, amely a memóriatárolás-
ban is szerepe van.

amelynek

Tegnap talált □ található egy kulcsot.

Egyszerű nyelvbotlás; téves szótalálás**Megakadás**

A másik tudtára adjam, amit tudni akarok.
 Akkor segítsetek már gyűj ☐ öö gyűjteni!
 Az inven ☐ az intervenció nagy.
 Azon gondoltam ☐ azon gondolkodtam, hogy ez
 miért van így.
 Csi ☐ csontleves lesz ma ebédre.
 Enyhén válogatott cit ☐ szitkozódások...
 ...értelmi fanyalgás című kategória értelmiségi
 fanyalgás című kategória.
 Huszonnyolcadikán meglátja a Flóra ☐ meglátogatja a Flóra.
 Kint van a kötényen ☐ kötélén.
 Mentek a szépszállítmányok.
 Nem lehet szállást fogni ☐ foglalni.
 Nemegyszer volt, hogy évi évödtünk otthon.
 Ötletgazdasága az páratlan.

Szándékolt közlés

amit tudatni akarok

szépszállítmányok

ötletgazdasága

Kontamináció; metatézis**Megakadás**

Ettől a kpétsör ☐ képtől sirtam el magam.

Kontamináció; téves szótalálás**Megakadás**

Ezzel is növelve a hangulatot, vagy legalábbis fokozva azt.

Szándékolt közlés

növelve a jókedvet × fokozva a hangulatot

Metatézis; morfológiai, szintaktikai hiba**Megakadás**

Nem lehetek ilyen bélpoklos! Hazabírom ki ☐ kibírom haza.

Szándékolt közlés

hazáig

Morfológiai, szintaktikai hiba; perszeveráció**Megakadás**

A fordításban és ezekben ☐ ben...
 A könyvet, a muzulmánok szent könyvet ☐ könyvét...
 Minden összeesküdi ☐ minden összeesküdik, na, összeesküszik ellenem.

Morfológiai, szintaktikai hiba; téves szótalálás

Megakadás

Amúgy is magamtól jön.
Tegnap ők látott bennünket.

Szándékolt közlés

magától
ő

Morfológiai, szintaktikai hiba; újraindítás

Megakadás

Nem lát ☐ láttad, az előbb mit csinált?

Perszeveráció; téves szótalálás

Megakadás

Az este felhőszaladás volt.
Ez idő alatt szövődött köztük valaki ☐ ööö ☐ valami.
Hamarosan elhangzik Beethoven nyitása ☐ ööö nyitánya.
Különben én már nem értek együtt ☐ egyet akkori önmagammal.
Mert hát minek tanuljon ő matekot vagy fizikát, ha az őt nem is érdekli.
Mind a nyolc meccset KO-val verte.
Nekem egyáltalán nincs miért félt ☐ ő ☐ fél-nem.
Természetes kivisz kivitelezés...

Szándékolt közlés

felhőszakadás

érdekli
nyerte

Téves szótalálás; változtatás

Megakadás

Nem fejez ööö illetve tovább, vagyis végig tartotta az óráát.

Újraindítás; változtatás

Megakadás

Én magam is ezt ☐ mind a ☐ mind a két utat bejártam.

A PEDAGÓGUSJELÖLTEK KAZINCZY-VERSENYE

A pedagógusjelöltek Kazinczy Ferencről elnevezett szép magyar beszéd versenyének 34. országos döntője

Pannon Egyetem, Veszprém, 2006. április 12–14.

A hazai és a határon túli felsőoktatási pedagógusképző intézmények hallgatói számára rendezett verseny tudományos ülésszakkal kezdődött a Veszprémi Akadémiai Bizottság székházában, amelyet a Pannon Egyetem Bölcsészettudományi Karának dékánja, dr. Földes Csaba egyetemi tanár és dr. Gósy Mária tanszékvezető egyetemi tanár (ELTE), a Kazinczy Ferenc szép magyar beszéd verseny elnöke nyitott meg. Ezt követően öt előadás hangzott el:

Adamikné Jászó Anna: *A főiskolai hallgatók beszéde;*

Jancsó Miklós: *A beszédtechnika oktatásának története az erdélyi magyar felsőoktatásban;*

Bóna Judit: *A beszédtempó;*

Imre Angéla: *A hangsúlyészlelés és a helyes hangsúly;*

Rozgonyiné Molnár Emma: *A beszédtechnika tanítása.*

A hallgatóság soraiban mind a résztvevők, mind a kísérő tanárok, valamint a helyi egyetemi hallgatók is jelen voltak.

A tudományos ülésszak után a Pannon Egyetem rektorhelyettese, dr. Kristóf János egyetemi tanár köszöntötte a résztvevőket a Bölcsészettudományi Kar fogadásán, amelyen dr. Mang Béla felsőoktatási helyettes államtitkár is részt vett és felszólalt: hangsúlyozta a verseny jelentőségét a pedagógusképző intézmények körében. Gósy Mária megköszönte a minisztérium támogatását, és megfogalmazta reményét, hogy a minisztérium a továbbiakban is segíti ennek a három és fél évtizedes nemes vetélkedésnek a folytatását.

A 2006-os Kazinczy-versenyen az előzőeknél jóval több felsőoktatási intézmény képviseltette magát: több mint 30 hazai és határon túli tanárképző főiskola, tanítóképző főiskola és egyetem 65 hallgatója és kísérő tanára vett részt a háromnapos rendezvényen.

A versenyen a kötelező és szabadon választott szövegek bemutatásáról hangfelvétel készült, amely tudományos kutatás alapjául szolgálhat.

Április 13-án reggel a szabadon választott szövegek előadásával kezdődött meg a viadal. A versenykiírásban szereplő szövegválasztási szempontokat (20–21. századi, esetleg 19. századi magyar publicisztika, esszé vagy kevésbé párbeszédes széppróza) a résztvevők többsége figyelembe vette. Mikszáth Kálmántól kezdve egészen Esterházy Péterig tallóztak a diákok, bár idén is, mint minden

évben, Kosztolányi Dezső és Márai Sándor művei voltak a legnépszerűbbek, illetve sokan választottak a (magyar) nyelvről szóló szöveget (pl. Nagy László, Illyés Gyula, Kodály Zoltán, Bárczi Géza tollából). Némelyik szöveg túl hosszú volt, jócskán meghaladta a versenyfelhívásban szereplő 3 percet, ezekben az esetekben a zsűri – a versenyzőknek az előző évekhez képest magas létszáma miatt – szigorúan vette a versenykiírást, és leállította az előadót. Néhány, különböző okok miatt nem megfelelő – pl. műfajilag nem ideillő (ismeretterjesztő jellegű) vagy nem felolvasásra, előadásra alkalmas, illetve párbeszédes – szöveget is hallhattunk.

A délután elhangzó kötelező szöveg egy eredetileg is kifejezetten felolvasásra szánt textus volt: a rövid részlet a veszprémi születésű nyelvésznek, Simonyi Zsigmondnak a nyelvtudományról megfogalmazott gondolatait tartalmazta. A Simonyi-gondolatok értő és értető bemutatásának a nehézségét elsősorban a szöveg nyelvezetének régiessége, 19. századi volta okozta a versenyzőknek.

A zsűri munkájában Gósy Mária elnöklete alatt Wacha Imre, Hoffmann István (Debreceni Egyetem), Nyirkos István (Debreceni Egyetem), N. Markó Julianna (Berzsenyi Dániel Főiskola), Süveges Gergő (Magyar Televízió), Markó Alexandra (ELTE), Fodor Eszter (Pannon Egyetem) és Pelczéder Katalin (Pannon Egyetem) vettek részt.

A zsűri feladata a versenyzők szokatlanul nagy száma (46) miatt nehéz és fárasztó volt, de a zsűritagok mindvégig egyetértésben és gördülékenyen dolgoztak, még ha volt is, aki szigorúbban, és olyan is, aki engedékenyebben pontozott. A reményeink szerint igazságos végeredmény éppen a zsűri sokszínűségének, változatos összetételének (egyetemi oktatók, tévébemondó, Kazinczy-díjas egyetemi hallgató) volt köszönhető.

A maximálisan elérhető 180 pontból az első helyezett 155 pontot, a 45. helyezett 68 pontot ért el. Ez azt mutatja, hogy a mezőny – felkészültségét tekintve – meglehetősen vegyes volt.

Az első 14 helyezett kapott Kazinczy-érmet, az ő pontszámaik között szinte csak 1-1 pont különbség jelentkezett.

A szervező intézmény büszke arra, hogy a nagyon fiatal (mindössze 5 éves) Magyar Nyelvtudományi Tanszék hallgatói közül kettő is a díjazottak közé kerülhetett.

A Kazinczy-érmes versenyzők sorrendben a következők: Egyedi Mónika (Pázmány Péter Katolikus Egyetem), Bak Anita Zsófia (ELTE BTK), Fenyvesi Zoltán (Pécsi Egyetem), Kéner Balázs (ELTE BTK), Szabó Pálma (Pannon Egyetem), Perczes Nikolett (ELTE Tanító- és Óvóképző Főiskolai Kar), Meleg Gábor (ELTE BTK), Nagy Róbert (Szegedi Tudományegyetem), Molnár Anikó (Pannon Egyetem), Dávid Ádám (Pázmány Péter Katolikus Egyetem), Deák Vera (ELTE BTK), Tulman Géza (Berzsenyi Dániel Főiskola), Hódi Viktória (Szegedi Tudományegyetem), Vanya Barbara (Konstantin Egyetem, Nyitra).

A díjazottak mellett minden versenyző értékes könyvjutalomban részesült az Oktatási Minisztériumnak, a Pannon Egyetemnek, a Veszprém Megyei Eötvös Könyvtárnak és Veszprém Megyei Jogú Városnak köszönhetően.

A 2007. évi 35. Kazinczy szép beszéd verseny szervezője az Esztergomi Tanítóképző Főiskola lesz.

S. Novotny Júlia – Pelczéder Katalin

Simonyi Zsigmond: A nyelvtudományról
(Felolvasás a veszprémi iparos ifjak képző és segélyző
körében 1877. január 2-án – részlet)

A verseny kötelező szövege

S ha itt előadást akarok tartani, miről szólanék egyébről, mint arról, amivel folyvást foglalkozom, a nyelvtudományról? Akik figyelemmel kísérték az itt tartott előadásokat, többnyire természeti, anyagi dolgokról hallottak, ritkábban az emberi elme működéseiről. Mert hamarabb és többször szokott az ember a külvilági, szemmel látható, kézzel fogható dolgokkal foglalkozni, mint saját elméjével s az ebben történő változatos eseményekkel. Pedig semmi sem lehet nagyobb befolyással szellemi haladásunkra, mintha magunkat, saját belső életünket mentől jobban megismerjük. S az emberi elmének egyik legcsodásabb műve a beszéd, a nyelv. Azt mondja egy hírneves író „Még nem mondhatjuk meg, mi a nyelv. Lehet a természet szüleménye, lehet az ember műve, vagy az Isten adománya. De akármelyik körbe tartozzék, a maga körében nincs, ami fölülhaladná, avagy csak hozzá fogható volna is. Ha a természet szüleménye, akkor a természetnek legutolsó és betetőző terméke, melyet csak magának az embernek tartott fenn. Ha az ember műve, akkor a művész-embert majdnem isteni teremtvé emeli. Ha isteni ajándék, akkor Istennek legnagyobb adománya, mert általa szólott Isten az emberhez, és általa szól ember az Istenhez, imádvá, könyörögve, elmélkedve.” A régi indiaiak szent énekeiben isteni rangra emelték a nyelvet, himnuszokat mondtak neki, s azt mondták róla, hogy az istennel volt együtt kezdettől fogva s hogy csodás dolgokat mívelt; bölcsaik pedig azt mondták, hogy a nyelv és a lehellet házasságából születik az emberi szellem. Az újabbkori tudósok közül is mondják némelyek, hogy a nyelv előbb megvolt, mint az ész, vagy legalább hogy együtt fejlődött az embernek esze és nyelve, hogy „nincs nyelv ész nélkül és nincs ész nyelv nélkül”, és általánosan elismerik, hogy a meghatározottabb válaszfalat ember és állat között éppen a nyelv képezi.

A 2006-os döntő kötelező szövegének értelmezése és előadásának egyfajta megközelítése

Az elmúlt három év Szép magyar beszéd országos döntőjének kötelező szövegei közül első ránézésre a 2006-os, a legutóbbi tűnik legnehezebbnek a téma tagolásának vizuális jelölése szempontjából. Magyarán mondván nincsenek mikroszerkezeti egységek, így a „faltól-falig” sorok átlátása, átfésülése nagyobb figyelmet kíván az előadótól. Ám, ha tüzetesebben megnézzük, a szöveg értő-értető előadását nem elsősorban a tipográfiája határozza meg. Sokkal inkább az, hogy szerzője melyiket szánta olvasásra, s melyiket megszólaltatásra. Vagyis, hogy az esszé, az értekező próza, a tudományos ismeretterjesztés műfajának írott vagy beszélt változata-e a kiválasztott alkotás.

Horváth János írása a Pázmány-életmű egy fontos darabját jellemzi, elemzi, hatását értékeli (2004). Németh László a kritika feladatáról értekezik vallomással, önmegszólító modorban (2005). Mindkettő gyönyörű nyelvezetű, hatalmas szófüzérekkel, szóképekkel, szerkezeti jelzőkkel, felsorolásokkal stb., de ezek a művészi eszközök ilyen töménységben nehezen befogadhatók még a vajt fülű hallgató számára is. Pláne, hogy a versenyzők mindössze 15-20 percet kapnak az értelmezésére.

Mennyiben más a 2006-os Simonyi Zsigmond-írás? Annyiban, hogy ez nem „írás”, hanem beszéd. Minden kétséget kizáróan maga a szerző mondja ezt ki az első mondatában. Egyértelmű, hogy egy szónoki beszéd – felolvasó ülésen való felszólalás – bizonyos részéről van szó. *S ha itt előadást akarok tartani* – utal arra, hogy mi csak bekapcsolódunk a szónoklatba, mert vélhetően előtte már elhangzott az ún. *captatio benevolentiae*. A nyitó mondat a téma megadása (*docilitas*), vagyis: én, a nyelvész miről szólhatnék, mint a nyelvtudományról. A figyelem megragadását (*attentio*) pedig azzal éri el Simonyi, hogy szembeállítja az addig elhangzottakkal az ő mondandójának lényegét: a szellemi haladás feltétele a nyelv megismerése. Majd konklúziómonddal zárja az első részt: *S az emberi elmének egyik legcsodásabb műve a beszéd, a nyelv*.

Az új gondolategység, a 2. rész, e tétel kifejtése egy idézettel. Ez amolyan érvelő-narratív rész hatásos gondolatritmussal, egy dialektikus érvsor, gradatív fokozás formájában: *Ha a természet [...] akkor [...] az embernek tartott fönn. Ha az ember műve [...] isteni teremtvé emeli. Ha isteni ajándék [...] akkor Istennek legnagyobb adománya*.

Ezt folytatja, fokozza a harmadik egységben emlegetve *az isteni rangra emelkedett nyelv*-et, valamint egy újabbkori tudós vélekedését chiazmusban összegezve: *nincs nyelv ész nélkül, és nincs ész nyelv nélkül*. Simonyi Zsigmond az elmondottak alapján arra a következtetésre készíti a hallgatóját, hogy *az ember lényege a nyelv, amely elkülöníti őt az állattól*. (Eddig tart a szónoklat számunkra, de nyilván folytatódik példák, érvek sorával...)

A verseny résztvevői kb. 80%-ban fölfedezték ezeket a gondolati egységeket. Akik nem, azok többnyire csak szószerkezeteket, mondatokat görgettek végig előadásukban. Nézzünk most egy hiteles interpretációs lehetőséget, és nézzük a kevésbé sikeres produkciók okait.

Mivel narratív, érvelő szövegről van szó, minden pátosztól és indulattól mentes a hangvétele. Zárójelben azért meg kell jegyezni, hogy a szerzőt és címet látva többen rögtön átszellemültek. Ennek egyik oka lehet az, hogy a középiskolai Szép magyar beszéd vetélkedőkön még mindig az édes anyanyelvünkről szóló – kissé már közhelyes, „pohos” – zengzeményeket ajánlják tanítványaiknak a jóakarató tanárok. Ezért kezdi a kezdő diák középhangsávjánál, normális hangfekvésénél egy regiszterrel magasabban, kenetteljesen, elrebegve ezt az amúgy tárgyilagos szöveget. Hasonló okból hiteltelen az ún. **túldramatizálás**, amikor a felolvasó elemelkedik a deszkákról, és olyan tragikus közegbe helyezi a szerző gondolatait, ahol azok sosem fordulhattak elő. Ugyanúgy, a **didaktikus, kioktató hangvételnél** is az a fő oka, hogy nem sikerült megérteni, miről kellene meggyőzni a hallgatóságot. Ilyenkor élnek gyakran a **puszta hangerőváltással**. Szövegritmus – tempó- és dallamváltás – nélkül pedig monoton, unalmas lesz a felolvasás. Amíg nem tudja megkülönböztetni valaki a **vezérszólamot** és az **aláfestéseket** egymástól, csak **írásjeleket hangosít meg agyontördelve a beszédművet**. (Akad elvéve olyan is, aki azért csinál a kétpercesből négyperces előadást, mert élvezzi a hangját: lebegtet, indokolatlan szüneteket tart, széttagol nagy ívű szólamokat.)

Ebben a szóközi beszédben több helyen is érdemes átértékelni a központosítást. E tekintetben kritikus máris az első mondat. Ha csupaszítom, így hangzana: *Miről szólnék másról, mint a nyelvtudományról!* Egy nyelvész szájából ez **evidenciakijelentés**. Nyilván csak formai jegy a kérdőjel, mivel kérdő névmás áll a mondat élén. Így nincs szükség kérdő mondati hanglejtésre, hanem ereszkedőre. Előfordult az is, hogy valaki az első tagmondatot már kérdésnek (főmondatnak) érezte, azaz a *tartani* szónál fölszöktette a dallamot. Ugyancsak felesleges a közeplő rész gondolatritmusának *Ha*-val kezdődő mellékmondatait kérdőre színezní. Ezek **feltevések és következtetések, nem kérdések**.

Az általam harmadiknak tekintett részben szintén érdemes racionalizálni a mondaton belüli írásjeleket.

1. állítás: *A régi indiaiak [...] a nyelvet. Himnuszokat [...] mívelt. Bölcsaik [...] szellem.*

2. állítás: *Az újabbkori [...] esze és nyelve: hogy [...] nélkül. És általánosan [...] képezi.*

Mindezt nem légzéstechnikai okokból, nem is szájbarágás céljából képzelem így, hanem az értelmi tagolás kívánja ezt.

Hadd essék szó még a **hangsúlyozásról**! Részben a rossz légvétel miatt csúszik el a hangsúly, de többnyire a rossz értelmezésből adódnak a tévesztések. Pedig például a 2. mondatban levő szembeállítás *többnyire ↔ ritkábban* szavai mellékhangsúlyosak, míg az utánuk következő jelzős szerkezetek főhangsúlyosak.

A középső rész anaphorája megadja alapját a főhangsúlyoknak. *Lehet a termé-
szet szüleménye, lehet az ember műve, vagy az Isten adománya.* Épp ilyen mó-
don hangoztatandó a részletező, kifejtő rész is a hangsúlyviszonyokat tekintve.

A 3. részben is van egy vitára okot adó mondat szerkezet. Ha így írta volna
Simonyi, hogy *A régi indiaiak szent énekeikben isteni rangra emelték a nyelvet,*
így az *indiaiak* alanya lenne a mondatnak. De itt az áll, hogy az *indiaiak szent*
énekeiben = szerkezetes jelző, és hozzá az *isteni rangra emelték* = igei állítmány
rejtett alannyal. (Vagyis ma így hangozhatna: *A régi indiaiak szent énekeiben is-
teni rangra emelkedett a nyelv.*) Ha tovább nézzük a mondat tömb állítmányait,
azok is mind rejtett alanyúak: emelték, mondták stb.

Érdekes, hogy a szövegben háromszor is szerepel a *magá*, *magunk*, *magának*.
Egyik esetben sem főhangsúlyos a személyes névmás nyomatékos alakja. Mégis
többeknél hallani lehetett így: *magának | az embernek tartott fönn.* Ha hangsú-
lyos, és még szünet is követi, akkor félreérthetővé válik a szándékolt gondolat.

Bár nem mondatfonetikai kérdés, inkább beszédtechnikai, mégis összefügg a
szövegértéssel a **koartikuláció**. Nem a szavakon belüli hangkapcsolati törvény-
szerűségekkel volt baj, sokkal inkább a **szóhatárokon** történő eltolódásokkal.
Aki blattol, vagyis – mint fentebb megjegyeztem – csak szó szerkezeteket, mon-
datokat próbál olvasni, annak nehéz az írásképtől elszabadulni, csak betűket
kapcsol össze.

A versenyt értékelvén egyetértettünk Süveges Gergő zsűritag társammal, hogy
a magyar nyelvben igen sok szó kezdődik *z*-vel, *t*-vel, *s*-sel. Mire alapozzuk ezt?
Nagyon gyakori a határozott névelő *z* hangjának **átrántása** a következő szóba,
pl. *a_zitt*, *a_zemberi*, *a_zújabbkori*, *a_zebben*, *a_zisten* stb. A *t-s* példák:
it_telőadást, *előadás_takarok*, *tartot_telőadásokat*, *szólot_tisten*, *közöt_téppen*.
Az *s* kötőszó: *s_éz*, *s_így*, *s_ott*, *s_ogy* stb.

Hogyan kerülhetjük el ezeket a hibákat? Ha a ritmus nagyon fontos elemére, a
szünetre is figyelünk. Látható, hogy az a **pillanatnyi értelmestítő szünet** hiány-
zott ezekből a szókapcsolatokból.

A kevés archaikus szóalak mechanikus felolvasása is vezetett tévesztésekhez.
Például *szólanék* helyett *szólnék*; *mentől* ~ *mentül*, *mentöl*; *mivel* ~ *müvelt*.
(Szemantikai archaizmus: *lehellet* ~ *lélek*.)

Folytathatnám az észrevételeket, de remélem, ennyivel is adtam valami képet
arról, hogyan értelmeztem és hogyan interpretálnám én a szöveget. Végezetül, a
versenyzők legalább 70-75%-a más-más módon, de hitelesen hangosította meg
ezt a szöveget.

N. Markó Julianna

A MAGYAR ÉS A NEMZETKÖZI TUDOMÁNYOS ÉLET ESEMÉNYEI

11. Nemzetközi Klinikai Fonetikai és Nyelvészeti Konferencia (XIth Conference of the International Clinical Phonetics and Linguistics Association)

Dubrovnik, 2006. május

A Nemzetközi Klinikai Fonetikai és Nyelvészeti Társaságot (ICPLA) 1991-ben hozták létre azzal a céllal, hogy erősítsék a beszéd és a nyelv zavarai foglalkozó fonetikai és nyelvészeti kutatásokat. A kezdeti elgondolások szerint ezzel egyúttal a kutatók és a gyakorlati szakemberek kapcsolatát is szorosabbá szándékoztak tenni. A nyolcvanas évek végén merült fel a Nemzetközi Fonetikai Társaság kongresszusán, hogy szervezettebb kereteket kellene teremteni azok számára, akik a beszédkutatás klinikai oldalaival foglalkoznak. Egy ilyen keretben az új eredmények közzlése, az információcsere, a vitalehetőség nagymértékben fejlesztheti a tudományterületet. Ezt követően az angol klinikai fonetikusok szerveztek egy összejövetelt, amelyen döntöttek egy olyan új társaság megalapításáról, amely összefogja a patológiás beszéd folyamatokkal foglalkozó szakembereket. Az első szimpózium megszervezése Cardiffban (Wales, Nagy-Britannia) adta a lehetőséget az ICPLA létrehozására. Ez az első szimpózium a klinikai fonetika kérdéseiről szólt; a későbbiek pedig már nemcsak a fonetikai, hanem általánosabb nyelvészeti kérdéseket is felvetettek. A huszadik század végének ez a fajta törekvése nagyon sok más tudományban is tapasztalható volt; világossá vált ugyanis, hogy bizonyos témák, problémák megközelítése csak interdiszciplinárisan lehet eredményes.

A Társaság két évente szervez szimpóziumot, amelyet különböző országokban rendeznek meg. A második összejövetelt Londonban (1992), a harmadikat Helsinkiben tartották (1993). A következő konferencia helyszíne a tengerentúli New Orleans volt (1994), majd München (1996), Nijmegen (1997), Montreal (1999), Edinburgh (2000), Hong Kong (2002) és Lafayette (2004) következett. Ebben az évben ismét Európában rendezték meg a találkozót, Dubrovnikban, Horvátországban. A Zágrábi Egyetem Fonetikai Tanszéke vállalta a szervezést, fő szervezői dr. Damir Horga és dr. Vesna Mildner professzorok voltak.

A konferencia 156 résztvevőt számlált, 26 országból, akik 84 szóbeli és 69 poszterelőadást mutattak be. Négy plenáris előadás a klinikai fonetika és nyelvészet talán legfontosabb kérdéseit elemezte. Az első napon Miloš Judaš (Horvátország) az agy és a nyelv, illetve a beszéd kapcsolatának legújabb eredmé-

neyeit foglalta össze *Development and evolution of language-related regions of the human cerebral cortex* címmel. A második napon Fiona Gibbon (Skócia) az elektropalatográfia (EPG) klinikai felhasználásának új területeiről beszélt (*Current perspectives on clinical application of EPG*). A harmadik napon Gósy Mária (Magyarország) gyermekek beszédészlelési stratégiáit tárgyalta különböző zavarok esetében (*Alternative organization of speech perception deficits in children*). Végül az utolsó napon Dana Boatman (USA) a centrális hallás ép és hibás folyamatait elemezte klinikai esetek bemutatásán keresztül; előadásának címe *Cortical auditory systems: Normal and impaired speech sound perception* volt.

A szekció- és poszterelőadások különféle kutatások, vizsgálatok, folyamatok eredményeit ismertették. Foglalkoztak a cochleáris implantáció következményeivel, a zöngéhang zavaaraival, patológiás beszédfolyamatokkal, hangképzési és nyelvi zavarokkal, afáziával és diszfáziával, a nyelv és a beszéd neurológiai kérdéseivel, a motoros kontrollal, beszédészleléssel és beszédmegértéssel. Sok kérdést és elismerést kapott Hoffmann Ildikó (Szeged), aki előadásában az Alzheimer-betegségben szenvedők beszédének tempóvizsgálatával foglalkozott. A szervezők 162 oldalas kötetben publikálták az előadások összefoglalóit, amelyek jól használhatók a későbbiekben is a témákban tájékozódni kívánók számára. A konferencia szervezése mintaszerű volt, amihez a tenger közelsége és Dubrovnik gyönyörű óvárosa is hozzájárult. A következő ICPLA-konferenciát két év múlva Törökországban rendezik meg.

Gósy Mária

5. nemzetközi konferencia a beszéd motoros kontrolljáról (5th International Conference on Speech Motor Control)

Nijmegen, 2006. június

1985 óta öt évente rendezik meg a hollandiai Nijmegenben az ezzel a témakörrel foglalkozó szakmai találkozót. A feladatot önként vállalta a nijmegeni Radboud Egyetem, azóta is fő szervezője az eseménysorozatnak, elnöke a kezdetektől dr. Ben Maassen. Már az első konferencia is sikeres volt, az azóta megrendezett valamennyi esemény hasonlóan magas színvonalon közvetítette ennek a gyorsan fejlődő területnek az eredményeit. A beszéd motoros ellenőrzése, kontrollja interdiszciplináris megközelítést igényel, akár az épen működő, akár a patológiás folyamatokról van szó. Ennek megfelelően a konferencia résztvevői fonetikusok, logopédusok, beszédklinikusok, neurológusok, agykutatók, pszichológusok, pszicholingvisták, neurolingvisták, audiológusok, gégszerek voltak. Az első konferencia (11 évvel ezelőtt) tárgya még főként a dadogás sajátosságainak megismerését célozta; az azt követő találkozók már egyre szélesebbé vál-

tak a tématerületek. Az előadások már általánosabban szóltak a motoros kontroll fejlődéséről, a magasabb szintű pszichomotoros és pszicholingvisztikai folyamatokról, a patológiás beszédet pedig kiterjesztették más zavart működésekre is. A harmadik konferencián előtérbe kerültek az agyi képalkotó eljárásokkal kapott eredmények bemutatásai; majd a negyedik alkalommal a magasabb szintű kognitív folyamatok és a beszédprodukciónak az összefüggéseire irányult a hangsúly. Ilyen előzmények után került sor az ötödik konferenciára ez év júniusában. Mintegy 140 résztvevő (25 ország képviselőjében, Braziliától Japánon, Tajvanon, Kínán, Ausztrálián át Amerikáig, Kanadáig és több európai országig) hallgatta meg az egy szekcióban szervezett előadásokat, és nézhette meg a posztereket a kijelölt időkben. Összesen 123 előadást fogadtak el (erőteljes szűrést követően). A kutatások egyre nagyobb mértékben mutatták az interdiszciplinaritás jellegzetességeit, aminek a hátterében az integráció mint kulskonceptió volt felfedezhető. Ez azt jelenti, hogy a kutatók az egyes témák modelljével és elveivel általánosságban, a beszéd hallási, vizuális és motoros megjelenésével pedig részleteiben foglalkoztak. Ezen belül akadt példa a biomechanikai, neurológiai megközelítésre, a genetikai tényező tekintetbe vételére, a beszéd motoros sajátosságainak és a tanulási folyamatoknak az összefüggéseire, illetve a különféle beszédzavarok sok szempontú elemzésére. A konferencia-előadások eredményei újra és újra igazolták, hogy a beszéd motoros kontrolljának kérdéseit kognitív, nyelvészeti, pszichológiai összefüggésben kell szemlélni, minden terület saját eszköztárának adatai alapján; a szintézishez azonban valamennyinek a tekintetbe vétele szükséges.

Az ez évi konferencia fő témái az alábbiak voltak: a beszéd motoros kontrolljának elmélete és modellálása, a beszéd neurológiai szerveződése, a beszédzavarok genetikai aspektusa, a nyelv- és beszédprodukciónak felelős neurális rendszerek, az ép beszéd és a beszédzavarok funkcionális agyi működésének elemzése, a beszéd motoros kontrolljának fejlődése, a dadogás, a neurogén beszédzavarok, klinikai fonetika különböző neurológiai állapotokban, klinikai vizsgálatok és terápiák, a beszéd motoros folyamatainak korszerű kutatási kérdései. Több előadás foglalkozott a dysarthria és az apraxia kérdéskörével, a genetikai eredetű beszédzavarok sajátosságaival, gyermek- és felnőttkori beszédprodukciónak a nehézségekkel. Rendkívül érdekesek voltak az ép beszéd folyamatok közvetlen artikulációs mozgásának eszközös (artikulográf) vizsgálatai különböző életkorokban. Megható volt az az előadás, amelyik a dadogás motoros problémájának agyi hátterét elemezte, mivel az előadó maga is e beszédzavarban szenvedett, és így kutatási eredményein túl saját személyes tapasztalatait is megosztotta a konferencia hallgatóságával.

Tekintettel a téma interdiszciplináris voltára, az előadók nagyon fegyelmezten tartották előadásait annak érdekében, hogy az adott területen kevésbé jártas résztvevő is megérthesse azokat. Ezt szolgálta a magas szintű szemléltetés (számítógépes kivetítő, film- és hangbejátszás) is. A szervezők külön gondot for-

dítottak arra, hogy legyen idő valamennyi poszterelőadás megtekintésére, és a szerzőkkel elmélyült megbeszélés folytatására.

Az előadók között magyar szerzők is voltak, előadásuk címe: *Disfluencies in the spontaneous speech of normally developed and SLI children* (Gósy Mária, Horváth Viktória, Csabai Katalin). Az előadások kivonatait a *Stem-, Spraak- en Taalpathologie* című folyóirat 14. számaként adták közre igényes kivitelben; az összes absztraktot ugyanazon logika szerint felépítve publikálták.

A szervezés mind tudományos, mind egyéb szempontból mintaszerű volt. Böven adódott idő a szakmai beszélgetésekre, vitákra, kérdésekre, válaszokra az előadásokat követően és a szabadidőben is. A szervezők gondoskodtak arról is, hogy a komoly munkát kellemes szabadidő váltsa fel esténként, például egy középkorú várban, ahol a kornak megfelelően öltözött harsonás fogadta a vacsorára érkezőket, amikor a várárok fölött a függőhídra léptek...

Gósy Mária

Az Amerikai Akusztikai Társaság (Acoustical Society of America) 151. konferenciája

Az Acoustical Society of America (ASA) 1929-es alapítása óta az akusztikával foglalkozó tudományos közösség meghatározó szervezete. Az ASA minden évben két konferenciát szervez, a szerző a 151-ediken vett részt 2006. június 5-től 9-ig. Ezt az Egyesült Államok legkisebb területű államában, Rhode Islandon rendezték meg. Az állam fővárosa és egyben a konferencia helyszíne Providence volt. Ez a város számos egyetemnek ad otthont, például a jó nevű Brown Universitynek.

143 előadás – 32 szóbeli és 111 poszter – foglalkozott a beszédtudomány különböző területeivel, ezen belül elsősorban a produkció, percepció, modellezés, elsajátítás és digitális feldolgozás aktuális kérdéseivel. Az összes előadás absztraktja megtekinthető a *The Journal of the Acoustical Society of America (JASA)* 2006. májusi számában (Vol. 119. No. 5. Pt. 2. 3207–3450). Itt négy érdekes előadást emelünk ki: kettőt a beszédpercepció, egyet a beszédprodukció és egyet a beszédpszintézis területéről.

John Kingston (University of Massachusetts) előadásában azt a hipotézist vizsgálta, hogy a beszédhangok azonosítására szolgáló egyes ismertetőjegyek nem egyformán fontosak a percepció folyamatában – például a zárhangok képzési helyének meghatározásában jobban hagyatkozunk a zárhangot követő magánhangzó formánsmeneteire, mint a zaj spektrumára. Carlos Gussenhoven (Radboud Universitat és University of London) a fizikai és nyelvi hangidőtartamok viszonyát vizsgálta: a veláris magánhangzók akusztikailag hosszabbak

(az állkapocssüllyesztés és -visszaemelés sok időt igényel), de a hallgatók nem érzik hosszabbnak, mert valószínűleg a feldolgozás során ellensúlyozzák ezt a jelenséget.

Steven Lulich (MIT) egy akusztikai modell számítógépes szimulációjával kimutatta az alsó légutak fonetikai szerepét. Az alsó légutak rezonanciái a formánsokhoz hasonlóan alakítják a zöngés hangok spektrumát, de a formánsokkal ellentétben nem erősítik a rezonanciafrekvencia körüli harmonikusokat, hanem gyengítik őket, valamint frekvenciájuk állandó. A második alsó légúti rezonancia természetes határként szolgál az elől és hátul képzett magánhangzók elkülönítéséhez: ha a második formáns frekvenciája magasabb, mint a második alsó légúti rezonancia, akkor elől képzett magánhangzóként érzékeljük, ha alacsonyabb, akkor hátul képzettként.

Susan Hertz (NovaSpeech és Cornell University) „hibrid” beszédszintetizátorával olyan tesztmondatokat állított elő, amelyekben a hangsúlyos magánhangzók egy bemondó természetes ejtéséből származnak, míg a többi hangot más beszédrészletekkel helyettesítette – egy másik bemondótól származó felvételdarabokkal vagy formánsszintézissel előállított szegmensekkel. Beszélőfelismerési kísérletek azt mutatták, hogy „hibrid” mondatokkal ugyanolyan könnyen felismerték a bemondót a kísérleti személyek, mint természetes felvételek esetén.

A konferencia előterében akusztikai kiállítást szerveztek. Ezen szerepelt egy videoismertető a Kempelen-féle beszélőgép rekonstruált változatáról, amelyet Olasz Gábor és Böhm Tamás készített (a beszélőgépet Olasz Gábor és Nikléczy Péter építette). A videót a résztvevők nagy érdeklődéssel fogadták, így megnyerte a kiállítás közönségdíját.

A konferenciát joggal nevezhetjük „diákbarátnak”. Diákok számára ingyenes volt a részvétel, a legjobb diákelőadások magas pénzjutalmat kaptak, és számos programot szerveztek annak érdekében, hogy a tanulók megismerkedhessenek egymással és a szakmai élet elismert személyiségeivel. Ennek érdekében például minden fiatal szerző kettesben együtt ebédelhetett egy, a területén dolgozó idősebb résztvevővel.

Böhm Tamás

BESZÉDKUTATÁS 2007

tudományos ülásszak

„Fonetika és pszicholingvisztika”

Az ülásszak ideje: 2007. november 15–16.

Helyszín: MTA Nyelvtudományi Intézet, földszinti előadóterem

Az előadás bejelentéséhez az előadó nevén, címén, e-mail és telefonadatain túl 10 soros összefoglaló is csatolandó. Jelentkezni és érdeklődni az alábbi elérhetőségeken lehet:

Horváth Viktória – e-mail: horviki@nytud.hu, telefon: 321-4830/172,
postai cím: Budapest VI., Benczúr u. 33., 1068

Jelentkezési határidő: 2007. május 31.

Az absztrakt alapján az előadás elfogadásáról 2007. június 30-ig küldünk értesítést. A lektorok által elfogadott kéziratok a *Beszédkutatás*-ban jelennek meg. A kéziratok leadási határideje minden évben szeptember 1.

A kötet szerzői

| Név | Munkahely | E-levél |
|----------------------|--|-----------------------------|
| Bata Sarolta | ELTE Fonetikai Tanszék | batasaci@freemail.hu |
| Bóna Judit | ELTE Fonetikai Tanszék | bona_judit@freemail.hu |
| Böhm Tamás | BME Távközlési és Médiainformatikai Tanszék | bohm@tmit.bme.hu |
| Gósy Mária | MTA Nyelvtudományi Intézet | gosy@nytud.hu |
| Gráczai Tekla Etelka | ELTE Fonetikai Tanszék | gracziketlaetelka@gmail.com |
| Horváth M. Szabolcs | Szent Rókus Kórház Neurológiai Osztály | |
| Horváth Viktória | ELTE Fonetikai Tanszék | horviki@nytud.hu |
| Imre Angéla | ELTE Fonetikai Tanszék | angelai@freemail.hu |
| Markó Alexandra | ELTE Fonetikai Tanszék | markoxa@ludens.elte.hu |
| N. Markó Julianna | Berzsenyi Dániel Főiskola | |
| Menyhárt Krisztina | MTA Nyelvtudományi Intézet | menyhart@nytud.hu |
| Mihajlik Péter | BME Távközlési és Médiainformatikai Tanszék | mihajlik@tmit.bme.hu |
| Németh Géza | BME Távközlési és Médiainformatikai Tanszék | nemeth@tmit.bme.hu |
| Nikléczy Péter | MTA Nyelvtudományi Intézet | nikleczy@nytud.hu |
| Olaszy Gábor | MTA Nyelvtudományi Intézet | olaszy@nytud.hu |
| Péteri Attila | ELTE Német Nyelvészeti Tanszék | peteri.attila@altavizsla.hu |
| S. Novotny Júlia | Pannon Egyetem Magyar Nyelvtudományi Tanszék | sino@vnet.hu |
| Zainkó Csaba | BME Távközlési és Médiainformatikai Tanszék | zainko@tmit.bme.hu |

